An Accuracy Controlled Combined Adaption-Optimization Scheme for Improving the Performance of 3D Microwave Devices over a Frequency Band

by

Dileep Nair

Department of Electrical and Computer Engineering
McGill University, Montreal
August 2008

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

Your file Votre référence
ISBN: 978-0-494-66589-3
Our file Notre référence
ISBN: 978-0-494-66589-3

Abstract

The design of 3D microwave devices can be improved by using computational optimization techniques combined with numerical simulations of the electromagnetic field. However, high accuracy field analysis is often computationally expensive and time consuming. One way to cut costs is to vary the accuracy level of the analysis at different stages of the optimization. This idea is based on the premise that the accuracy need not be constant throughout the optimization, and so the numerical analysis can be run more cheaply without compromising design quality.

This thesis presents a software system that minimizes the return loss of 3D microwave devices over a frequency band efficiently through accuracy control. It combines a custom gradient-based optimizer with a *p*-adaptive frequency-domain finite element solver. The solver computes the cost function and its gradient to a specified accuracy in a cost efficient manner. The *p*-adaptive solver comprises of two original components: an *a-posteriori* error estimator to evaluate the error in the cost function gradient, and an error indicator to identify the high error regions in the mesh. The optimizer controls the accuracy of the cost function evaluation through a link with the solver, specifying the required relative error for the gradient at each optimization step.

The combined adaption-optimization scheme was applied to 3D rectangular waveguide problems for validation: an E-plane miter bend, a U-bend, an impedance transformer and a compensated magic-T. For comparison, all the problems were also optimized using high-order finite elements at every step. Test results prove the computational efficiency of the new combined scheme at various stages of the optimization. In the early stages, when the element orders are low, the scheme is able to attain similar cost function reductions as the high-order analysis, with computational savings up to a factor of 25. Even in the late stages, when the accuracy is more stringent, the scheme manages a reduction in cumulative computation time of at least a factor of 4.

Abrégé

La conception de dispositif hyperfréquence 3D peut être améliorée par l'utilisation de techniques d'optimisations algorithmiques jumelées à des simulations numériques du champ électromagnétique. Cependant, la modélisation du champ électromagnétique à un niveau de précision élevé est très coûteuse en terme de nombres de calculs et de temps. Le coût peut être réduit en variant le niveau de précision utilisé à chaque étapes du cycle d'optimisation. Le concept est basé sur l'idée que le niveau de précision peut être varié pendant l'optimisation. Le résultat est que l'analyse numérique prendra moins de ressources en évitant une baisse de qualité.

Cette thèse présente un système informatique qui minimise l'affaiblissement de réflexion de périphériques micro-ondes 3D. L'éfficacité du système est augmentée en controllant de façon dynamique la précision des calculs. Le système combine un optimiseur basé sur les calculs de gradients avec un programme de calcul d'éléments finis de domaine fréquentiel p-adaptif. Le programme p-adaptive évalu la fonction du coût et son gradient à une précision fixée par l'optimiseur, et il a deux componsantes. Le premier est un estimateur d'erreur a-posteriori qui évalu l'erreur du gradient de la fonction du coût. Le deuxième est un indicateur qui identifit les zones d'erreurs importantes dans la maille. L'optimiseur contrôle la précision de l'évaluation du fonction du coût gràce a un lien avec le programme p-adaptive ou l'erreur relative du gradient est spécifiée à chaque étape de l'optimisation.

La methode d'optimisation adaptive a été utilisée pour valider plusieurs guides d'ondes rectangulaires 3D, donc un «E-plane miter bend», un «U-bend», un transformateur d'impédance, et un té magique compensé. La comparaison de tout les guides à été accompli par la modélisation par éléments finis de haute précision. L'éfficacité de la methode est supportée par les resultats obtenus à chaque étape de l'optimisation. Pendant les premières étapes, la précision des calculs est maintenue à un niveau plus bas. Les résultats sont comparables à celles obtenues par les analyses de haute précision, mais avec une diminution jusqu'à un facteur de 25 du coût des calculs informatiques. Même vers la fin de l'optimisation, quand la précision est plus importante et donc plus élevée, une réduction du coût des calculs par un facteur minimum de 4 est observée.

Acknowledgements

I would like to express my most sincere respect and gratitude to my supervisor, Professor Jon P. Webb, for his guidance throughout my research. This work would not have been possible had it not been for his kind support and patience throughout its duration. His considerable knowledge and keen insights into the various topics related to this research have been invaluable in advancing this project. I am very grateful for having had him as my advisor.

I would also like to thank the other professors of the Computational Electromagnetic group at McGill, Professors Dennis Giannacopoulos, David Lowther, Steve McFee, and Milica Popovic for their support and words of wisdom. Special thanks to my fellow research group members and acquaintances that I have come to know during the past few years at McGill. These people have not only been helpful to me with their research related ideas and discussions, but also with their friendship and camaraderie that made my time at McGill an enjoyable and enriching experience.

No words can express my immense gratitude to my parents, Mr. Gopal Nair and Mrs. Rugmini Nair, for their support and constant faith in me. I would especially like to thank my mother, who was always there to encourage and emotionally support me through the rough patches over the course of my research. I am grateful to my father for his constant encouragement that urged me forward, as well as for his advice that got me started along the research career path. I would also like to thank my brothers Vinod and Ramesh for their support throughout. I do not know if I would ever be able to repay the debt owed to my family for their patience, unconditional love and support through all these years.

Table of Contents

Table of Figures

Acronyms and Conventions

Acronyms

AVM	-	Adjoint Variable Method
AWE	-	Asymptotic Waveform Evaluation
CAD	-	Computer Aided Design
CFE	-	Cost Function Evaluation
DOF	-	Degree(s) Of Freedom
EM	-	Electromagnetic
FD	-	Finite Difference
FE/FEM	-	Finite Element Method
*N*D	-	*N*-Dimensional
PEC	-	Perfect Electric Conductor
PMC	-	Perfect Magnetic Conductor
VB	-	Visual Basic

Conventions

[A]	is a matrix where A_{ij} is the ij^{th} entry
{A}	is a column matrix where A_i is the i^{th} entry
A	is a 3D vector
$\{g^{(k)}\}$	is vector of geometric parameters at the k^{th} iteration of the optimization

1. Introduction

Due to the tremendous growth in computer technology over the past two decades, optimization techniques are now being used to find solutions to larger, more complex problems in varied fields like mathematics, applied sciences, engineering, economics, medicine, statistics, and so on [1]. In the field of high frequency electromagnetics (EM), this breakthrough in computing processing power has been of significant help in the accurate analysis and design of complex three-dimensional (3D) microwave devices, with designers relying more heavily on EM-based computer aided design (CAD) techniques than ever before.

Design optimization and accurate analysis of 3D microwave passive devices are required to meet the increasing demands for higher quality microwave components for wireless and satellite based communications, as well as for radar applications. Passive microwave devices such as power dividers, resonators, filters, waveguide transformers, etc., are critical components for microwave systems that are used extensively in these fields. Due to high industrial demands, there is more emphasis on reducing design time and costs, while yielding optimized designs for these devices, and CAD techniques offer a cost-efficient way to accomplish this. Yet, despite the advances in computing power, the EM analysis of complicated 3D problems is still an extremely time consuming process. This requires designers to devise new computational techniques that produce reliable designs at reduced computational costs.

A practical approach to efficient design is to generate an initial model using approximate, but fast, synthesis methods, and then as a second step, fine-tune this model using computational techniques to generate a design with improved performance. The initial design can be obtained using data from a design manual, or an approximate analytical formula, perhaps based on circuit theory [2], or by means of a knowledge-based approach [3]. The fine-tuning process can then be conducted by identifying the critical design parameters of the device and varying them to satisfy the performance targets. Computational optimization techniques can be used to model the tuning process, with the user-targeted performance parameter designated as the cost function that can be either minimized or maximized to satisfy the performance criteria. The type of optimization technique chosen will also have a significant impact on the overall cost efficiency of the fine-tuning process, as explained below.

The two main types of optimization techniques are deterministic and stochastic. Of the two, stochastic methods are attractive due to their ability to find the global optimum in a given design space populated with numerous local optima. However, they are computationally expensive, as numerous cost function evaluations (CFE) are necessary to effectively narrow down the selection to a set of good optimum points, and eventually obtain the proper global optimum. Moreover, for the fine-tuning case where the objective is to merely seek one of the numerous local optima lying near to the current design point, the application of stochastic methods is excessive from a computational point of view. Hence, the deterministic approach using gradient based optimization techniques is more practical in this context.

Gradient based direct optimizers make use of the sensitivities, or rates of change, of the cost function with respect to the design parameters. Such optimizers incorporate EM-solvers to compute the response values and their sensitivities for every optimization step. Extensive work in the past two decades has improved the efficiency of EM-solvers, and techniques based on the Finite Element Method (FEM), Integral Equation method, Finite Difference (FD) methods, Mode Matching methods and Method of Moments, as well as various hybrid techniques that combine these different individual methods, have been developed to handle complex 3D problems [4],[5]. Traditional EM-based optimization techniques use the finite difference approach to find the sensitivities, where the EM-solver is repeatedly invoked for perturbed values of the design variables. However, this is an inefficient approach, especially for problems with a large number of design parameters: at a given design point for a problem with n design parameters, $n+1$ cost function evaluations are needed to obtain the design response and its sensitivities. The overall computational cost becomes prohibitive for complex 3D problems, where each CFE is itself expensive.

As such, the computational costs involved in the analysis need to be cut by making the EM-simulator produce the cost function and its sensitivities cheaply. One efficient approach is to use the adjoint variable method (AVM) in conjunction with the

FE method, which has been previously applied to microwave field analysis [6]. This approach allows the calculation of the gradient value using only one extra simulation, regardless of the number of design parameters in the given problem [7], [8]. However, an alternate FE formulation presented in [9] takes this concept a step further by calculating the sensitivities at no extra cost beyond what is required for the regular solution, irrespective of the number of design parameters. The resulting computational savings for an iterative 3D analysis, such as the one used for the current optimization scheme, is substantial.

Further cost reductions can be achieved by introducing a feedback loop from the optimizer to the EM-solver as a means of controlling the cost function accuracy. The underlying idea is that the accuracy requirements associated with the cost function and sensitivity calculations do not necessarily need to be constant throughout: the accuracy at the initial stages of the optimization need not be as stringent as that required towards the end, when the process gets closer to the optimum. Providing a link from the optimizer facilitates the transfer of this information to the EM-solver and enables it to conduct the solution process in a more computationally prudent manner. This idea was pursued in [10] for the FE analysis of 2D microwave problems.

The accuracy link further ensures that the solution at each analysis step is numerically valid. If the EM-solver is unable to attain the accuracy level set for the given analysis point, it should communicate this to the optimizer, and the entire optimization-EM analysis process must be terminated, as any improvement beyond this point are meaningless from a numerical perspective. Consequently, the integrity of the fine-tuned optimal design produced is assured by the system. This is in contrast to normal practice, where the optimization is seeking the best possible optimum without regard for the accuracy of the numerical solution.

A schematic roughly illustrating the overall optimization scheme combining the optimizer module and the EM-solver module, as well as the role of the accuracy link in the process is given in Figure 1.1.

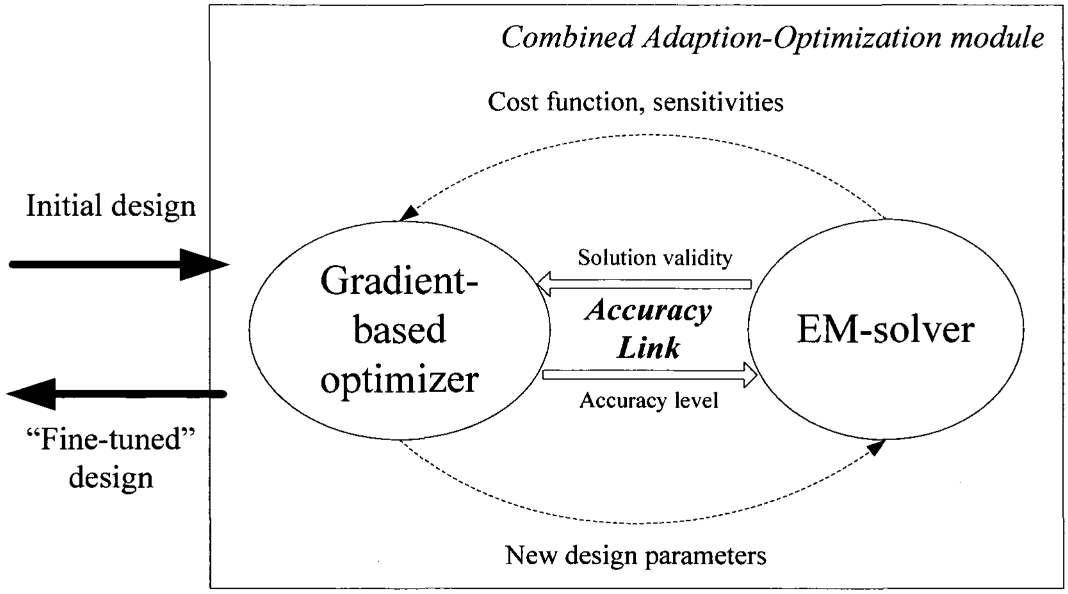

Figure 1.1: Schematic of the optimization scheme with accuracy link

For this work, the p-adaptive FE technique is chosen as the EM-solver method used by the optimization process. FEM is a well-established and versatile technique for high frequency EM analysis, capable of conducting accurate and efficient full-wave analysis of arbitrary complex 3D structures, aided by a well refined mesh that models the given geometry. The p-adaptive version of the FEM assigns the additional degrees of freedom by varying the element orders of the mesh. Accuracy control on the FE analysis through the link allows a more intelligent approach to the distribution of the additional degrees of freedom, thus improving the computational efficiency of the optimization process. An error analysis formulation for 3D problems within the p-adaptive FE formulation is used in conjunction with the accuracy control scheme to attain this. The fact that the FEM formulation can provide the sensitivities at no extra cost, as mentioned previously, contributes to the efficiency as well.

The objective of this research work is to implement an efficient optimization scheme combined with an EM-based solver using the ideas that are outlined in this chapter. The detailed description of the main components used in this scheme and how the different modules interact with each other to maintain that efficiency constitutes the remainder of the thesis.

Chapter 2 deals with the choice of optimization method used for the current analysis scheme. The advantages of using a gradient based system and how the gradient information can be used for computational cost reductions are illustrated in this chapter.

Chapter 3 introduces the numerical analysis module using FE, as well as the theory to compute the cost function and its gradients cheaply over a frequency range of interest. These are then used in a gradient-based technique that incorporates the changes introduced in Chapter 2.

Chapter 4 delves into the details of the *p*-adaptive FE formulation used for computing the cost function and their gradients inexpensively. The development of the error analysis scheme and the hierarchal elements that are critical for an effective *p*-adaptive method are outlined in this chapter.

The accuracy controlled optimization scheme linking the FE analysis and the optimization technique is covered in Chapter 5. Here, the role of the accuracy link connecting the optimizer to the FE analysis module in enforcing computational efficiency of the analysis is described in detail.

2. Optimization

The purpose of optimization is to arrive at the best possible decision in any circumstance, with the ultimate goal of improving the targeted characteristics of a given problem. When included in the design process, the optimizer guides a search in solution space in its attempt to improve a given design, and possibly obtain more innovative configurations [11]. Optimization methods have long been used to aid the design of electromagnetic devices [12], with applications in the field of microwave circuits and antenna design [13], [14]. The importance of optimization techniques in EM-design has been recognized by the CAD industry, and has incorporated optimizers into commercial simulation software packages [15], [16].

In general, optimization techniques can be divided into two groups: stochastic and deterministic optimization. Evolution strategies (ES) [17], genetic algorithms (GA) [18], [19], simulated annealing (SA) [20], artificial neural networks (ANN) [21], particle swarm optimization (PSO) [22] and Taguchi's method [23] are some of the stochastic algorithms that have been developed over the years. A review of the application of stochastic methods in the optimization of electromagnetic problems is given in [24]. The advantage of these methods is their ability to find the global optimum in a given solution domain. In addition, they are mostly independent of the search domain, thus making them useful when dealing with new problems in which the nature of the solution space is relatively unknown, or ill-behaved. The downside is the high computational cost associated with multiple design points scattered throughout the solution domain during its search for the global optimum [3]. This makes these methods considerably expensive for iterative 3D EM analysis.

Deterministic methods provide a more directed search by taking advantage of the local search space characteristics (like gradient information) in the vicinity of a design point in order to converge to a local optimum as quickly as possible. For most microwave design problems, a decent initial design is already available through fast synthesis methods, and bringing the solution to a nearby local optimum is often all that is required to produce a good design [25]. Furthermore, this "fine-tuning" can be accomplished at a much lower cost than stochastic methods. Hence, deterministic methods are better suited for implementing a combined EM-analysis-optimization scheme for the cost-efficient design of microwave devices.

The space within which the optimizer searches for an optimum is defined in terms of the set of design parameters that control the performance characteristics of the given problem. The design space is an N-dimensional Cartesian space with each co-ordinate axis representing a design variable g_i (i=1, 2,...N), and any solution to the optimization problem is a design point in this N-dimensional space.

For the current research, the optimizer is used to find the *minimum* of the given cost function; hence any reference to a local optimum in design space in the following discussions actually implies a local minimum.

2.1 Steepest Descent Method

Gradient based methods belong to the class of deterministic optimization techniques that require not only the cost function value but also its derivative. The gradient represents the rate of change of the cost function with respect to a set of design parameters (e.g. geometric dimensions) that define the problem. Since the negative of the gradient vector gives the direction of steepest descent, setting the search direction equal to that value will give the path direction along which the cost function decreases the quickest at a given design point [1]. This is the underlying principle of the Steepest Descent method.

Hence, given the cost function C and its gradient ∇C [1], the search direction $\{p\}$ is set as:

$$\{p\} = -\nabla C \tag{2.1}$$

where the gradient is comprised of values $\{\partial C/\partial g_1, \partial C/\partial g_2, ..., \partial C/\partial g_n\}$, and $\{g\}$ is the design parameter vector.

[1] The cost function gradient column vector $\{\nabla C\}$ will be denoted by ∇C throughout this thesis

2.2 Line Search

Once the search direction is set, a local minimum of the cost function along it is determined by an iterative process, which yields a series of design points:

$$\{g^{(k)}\} = \{g^{Start}\} + v^{(k)}\{p\} \tag{2.2}$$

where $\{g^{(k)}\}$ is a column vector of design parameters at the k^{th} step of the line search, $\{g^{Start}\}$ represents the initial design point that is used to calculate all subsequent design points on that line search, $\{p\}$ is the column vector representing the line search direction, and $v^{(k)}$ is the scalar that controls the step length along the line.

Finding the step length v that will minimize the cost function value along direction $\{p\}$ efficiently is of great importance. Several one-dimensional minimization methods have been developed to determine the step length for multi-variable optimization methods: elimination methods (Fibonacci and Golden section), which do not require gradient values, and interpolation methods (quadratic and cubic interpolation) that do use the gradient, are examples. For the current scenario, due to the availability of the gradient information, the more effective Davidon's cubic interpolation technique is implemented [26].

The aim of the one-dimensional minimization method is to find a non-negative value of v for which the function:

$$C(v) = C(\{g^{Start}\} + v\{p\}) \tag{2.3}$$

attains a local minimum, where $C(v)$ is the original cost function expressed in terms of the step length parameter alone (hence the designation one-dimensional optimization). Davidon's cubic interpolation method makes use of the first derivative information of the cost function to find the directional derivatives along the fixed search direction, given by:

$$G = \nabla C^T \{p\} \tag{2.4}$$

The directional derivatives are used to determine the upper and lower limits of a search interval within which the local optimum exists. Once set, the non-linear cost function is approximated using a cubic polynomial with respect to the parameter v within this interval, whose solution is an approximation to the ideal value of v that minimizes the cost function locally. An iterative process in which the initial interval is reduced for every subsequent attempt at finding a better approximation to the optimum is carried out until the solution satisfies certain termination criteria.

The implementation details of the cubic interpolation method are given below.

2.2.1 Search interval bounds

The objective is to establish the line search interval by finding two points along the search direction that satisfy the necessary conditions for a minimum value to lie in between.

Given an initial point represented by the design parameter vector $\{g\}$, let v_1 be a given value of v that corresponds to the design point, with directional derivative value $G_1 < 0$ and cost function value C_1 making v_1 a lower bound for the search interval.

To find the upper limit v_2 ($> v_1$), the following equation is used:

$$v_2 = v_1 + \frac{2(C_e - C_1)}{G_1} \tag{2.5}$$

where C_e is a preliminary estimate of the cost function's minimum value. In order to set the design point $\{g\}$ as the lower bound, the value of v_1=0 is used. Once v_2 is obtained, the corresponding values of C_2 and G_2 are calculated. Two possible scenarios that allow for a minimum to lie between v_1 and v_2 are shown in Figure 2.1.

In the first scenario, the directional derivative G_2 will have the opposite sign of G_1 (i.e., $G_2 > 0$), while the cost function values satisfy either condition: $C_2 > C_1$ or $C_2 < C_1$. The second scenario arises when $C_2 > C_1$, and with G_1 and G_2 having the same sign, v_1 and v_2 can still form the interval bounds within which a minimum can exist. If neither of these conditions is satisfied, the upper bound search will continue, with the current value v_2 doubled to obtain the next bound.

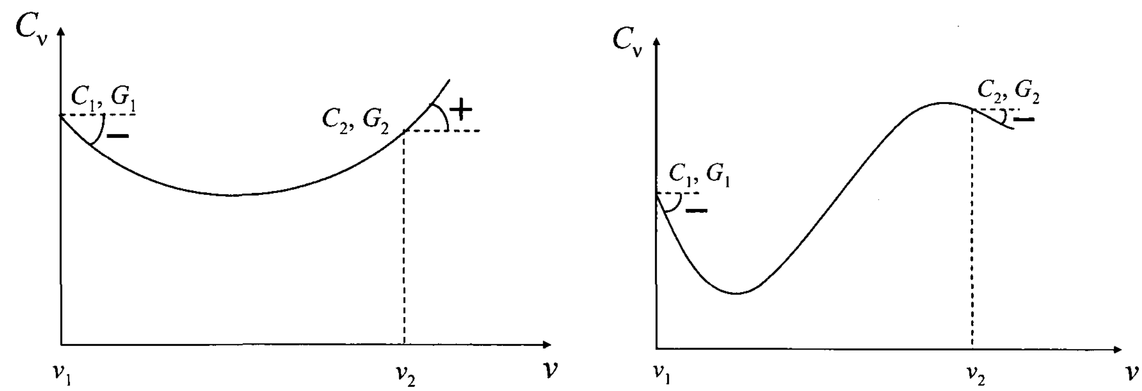

Figure 2.1 Line search interval bounds

2.2.2 Approximating the local minimum

Once the interval bounds are set, the cost function is approximated inside the interval using a cubic polynomial, whose solution provides an approximation to the local minimum.

Provided that the cost function values and the directional derivative values are known at both upper (v_2, C_2, G_2) and lower (v_1, C_1, G_1) bounding points, the following equation provides the minimum of the cubic polynomial fitting these values:

$$v^* = v_1 + (v_2 - v_1)\left\{1 - \frac{G_2 + w - z}{G_1 - G_2 + 2w}\right\} \tag{2.6}$$

where

$$w = \sqrt{z^2 - G_1 G_2} \tag{2.7}$$

and

$$z = \frac{3}{(v_2 - v_1)}(C_1 - C_2) + G_1 + G_2 \tag{2.8}$$

Equations (2.6)-(2.8) are used iteratively for finding a better approximation v^* for the minimum. If the current value of v^* is not a good enough approximation, the line search will continue with the value of v^* set as either the upper (v_2) or lower (v_1) bound of the new search interval. This assignment is based on selecting the interval within which the minimum could exist, as illustrated by two simple cases in Figure 2.2.

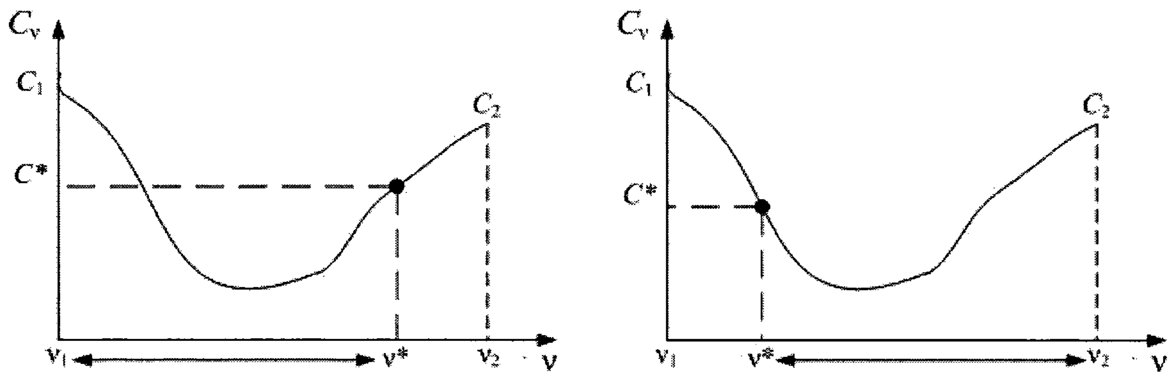

Figure 2.2 Possibilities for setting the new line search interval bounds

In the first case, with $v^* < v_2$, $C^* < C_2$, and with G_1 (< 0) and G^*(> 0) having opposite signs, (v^*, C^*, G^*) can be set as the new upper bound, replacing (v_2, C_2, G_2). For the second case, with $v^* > v_1$, $C^* < C_1$ and G^*(< 0) and G_2 (> 0), the points (v^*, C^*, G^*) and (v_2, C_2, G_2) will bound an interval with a minimum located within. This sequence is continued until the proper line search termination criteria are satisfied.

2.3 Line search termination

A variety of criteria can be utilized to terminate the line search process, once the approximation to the optimum is deemed accurate enough. Possibilities include:

(a) A design-parameter value based termination criterion, where the line search is terminated if the difference between successive design parameter values is below a certain threshold

(b) A cost-function value based criterion, where the line search is terminated if the difference between consecutive values of the cost function is below a certain threshold

A different approach is used in the present work, which takes advantage of the availability of the cost function gradient value for every analysis point.

The directional derivative, given in (2.4), indicates that at a local minimum along the current search direction, the vector quantities ∇C and $\{p\}$ become orthogonal to each other. The angle θ between them can be obtained by re-arranging the definition of the dot product:

$$\cos\theta = \frac{\|\nabla C\| \cdot \|p\|}{\nabla C^T \{p\}}. \tag{2.9}$$

where $\|\nabla C\|$, $\|p\|$ are the magnitudes of the vector quantities. Since the gradient is available at every analysis step, the angle can be found easily to determine its deviation from the current search direction. If it deviates by a significant amount, this indicates the possibility of another, better, minimum point lying in a different direction (indicated by the new gradient) than the current one. By judiciously selecting a threshold value for the angle θ in (2.9), the optimizer can terminate the current line search without having to wait to reach the local minimum, and launch a new one using the gradient as the search direction as soon as the threshold value is exceeded, thus avoiding the extra cost function evaluations (CFEs).

2.4 Implementation of the steepest-descent algorithm with cubic line search

A rough outline of the modified steepest-descent algorithm with the threshold angle termination criteria can be given as follows:

0. Set threshold angle θ_{th}; and initial point in design space
1. Compute C, ∇C for the design configuration.
2. If the current point is the initial one from Step 0, set the line search direction to $-\nabla C$; go to 5.
3. If the angle between the new $-\nabla C$ (obtained from the FE analysis) and the current line search direction exceeds θ_{th}, go to 4; else go to 5.
4. Start a new line search with search direction set to $-\nabla C$.
5. Find the new design point along the current line; go to 1.

A numerical analysis module approximates the $C, \nabla C$ values corresponding to each design point $\{g\}$ specified by the optimizer at various stages of the line search algorithm. The module discretizes the given geometry and numerically approximates the cost function and its gradient. Upon receiving these values, the optimizer then uses them for further processing depending on the line search stage (e.g., bounding point conditions, or the angle criteria for line search termination) to take necessary action for advancing the optimization analysis. Figure 2.3 gives a more detailed illustration of the line search algorithm, with the various stages connected to the numerical analysis module.

The initial geometry used for the pre-optimization analysis yields ∇C , which is designated as the initial search direction $\{p\}$. The design parameter vector $\{g^{Start}\}$, from which subsequent design parameter values for the given line search are calculated, is set using this initial value. Once the search direction and the lower bound are set, the next step is to determine an upper bound for the search interval, as explained in Section 2.2.1. The value for $\{g^{bound}\}$ to determine this is calculated using (2.5) and then (2.2), and the corresponding $C, \nabla C$ values that satisfy the upper bound conditions (as explained in Section 2.2.1) are sought. Section 2.2.2 explained how the line search step v^* is found for the cubic point analysis once the interval is set. The vector $\{g\}$ for the new design point lying in the direction $\{p\}$ is obtained using (2.2), and the corresponding values for $C, \nabla C$ are calculated. The line search termination condition based on the angle criterion is evaluated using (2.9). If the angle criterion is satisfied, the line search is broken, and the ∇C value becomes the search direction for the new line search. The vector $\{g^{Start}\}$ is updated using the last valid design point from the previous line search, which will also act as the lower bound (start point) of the new line search. For the scenario where the angle criterion is not satisfied, v^* is set as one of the bounds of the new reduced line search interval, and the process to find a better approximation to the minimum continues.

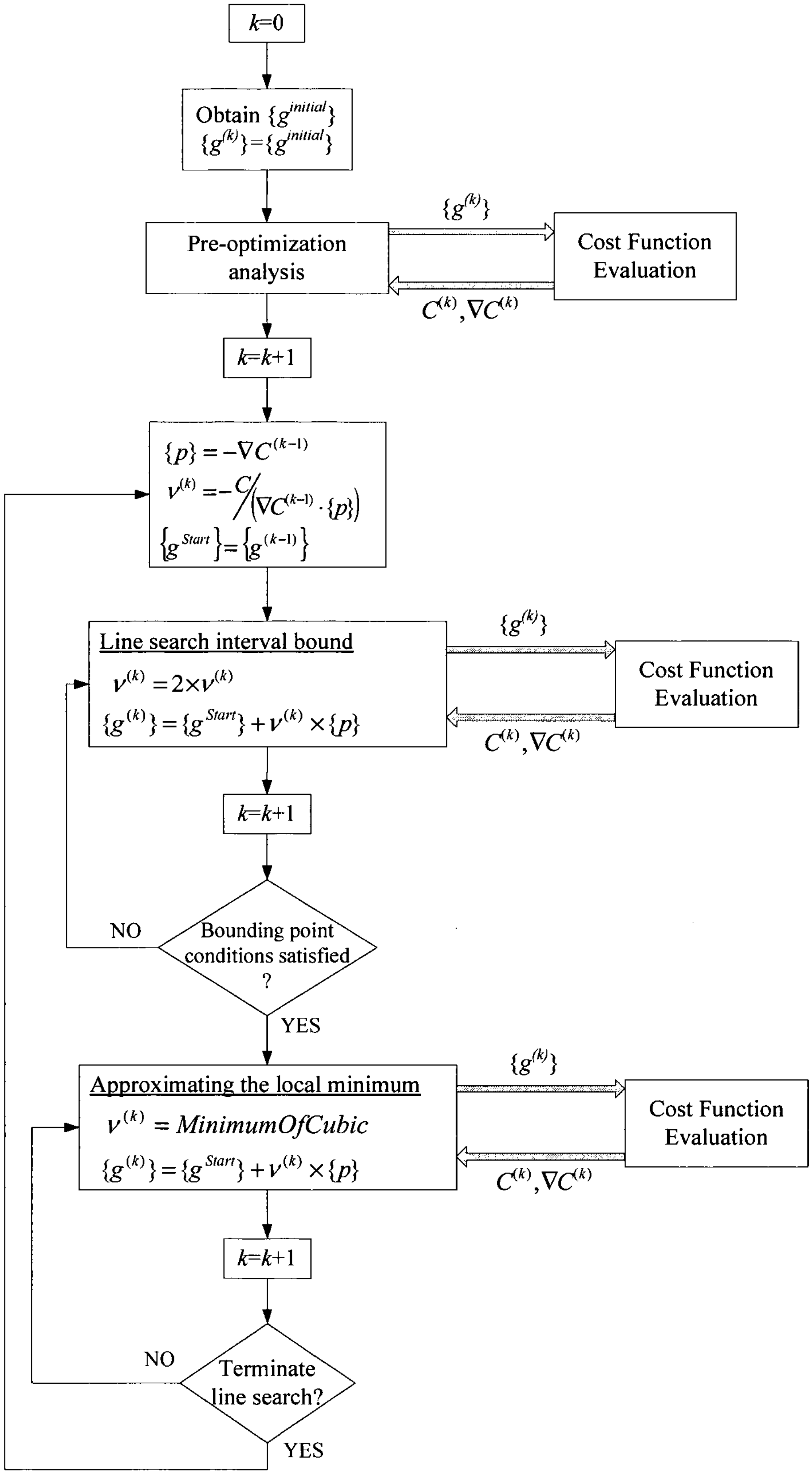

Figure 2.3 Implementation of the line search algorithm combined with the CFE module

2.5 Role of the optimizer in maintaining computational efficiency and reliability

In light of the increasing complexity of the problems dealt with by present-day designers, maintaining computational efficiency is a major concern, despite the considerable advances in computing power. One way of doing this is for the optimizer to take an active role in implementing checks and controls that help regulate efficiency, while maintaining reliability. This is in contrast to the usual passive role the optimizer plays in the traditional design process.

One attempt at reducing overall computational cost stems from the idea that the accuracy requirement for the solution in the initial stages of the optimization need not be as stringent as that required when close to the minimum. The cost function gradient value decreases from the first optimization step onwards, and the optimizer can set the accuracy requirement for a given line search using this as a benchmark, forcing the accuracy to increase as the optimization converges to the minimum. This results in significant computational savings for the optimization process, especially in the initial stages.

Once the accuracy requirements for a given line search point is set by the optimizer, the EM-analysis module must produce results that satisfy them. If unable to do so, the optimization process becomes invalid, as the numerical approximation to the cost function is no longer close enough to the true solution. Any improvement beyond this point is meaningless, and the optimizer should terminate the entire process. The last obtained valid solution becomes the best approximation to the local optimum that the optimizer can produce. This way, the optimizer is able to guarantee the validity of the solution, and consequently, the designer can have confidence in the integrity of the optimal design produced by the system.

Further computational cost savings for the optimization process can be obtained by calculating the cost function and its gradient for cheap at every analysis step along the way. How this is accomplished using the Finite Element method, as well as how the optimizer-controlled accuracy requirements are enforced on the FE analysis process, forms the basis of the remainder of this thesis.

3. Calculating the cost function and its gradient

To improve the device performance through optimization, the cost function must be formulated in terms of quantities that characterize the device. A common design goal for a microwave designer is to minimize the wave reflection at a port of a given device to attain maximum power transfer over the desired operating band. For this thesis, the cost function is formulated in terms of the return loss and minimized. In order to approximate the electromagnetic fields that are required to compute the cost function, an accurate field-based analysis tool capable of handling arbitrarily-shaped devices is necessary.

The Finite Element (FE) method is one of the most versatile and powerful analysis techniques used in computational electromagnetics due to its ability to handle problems with complex geometries. It is applied to a wide range of problems such as static, quasi-static, wave and transient systems, and problems with different material properties (non-linear, inhomogeneous and anisotropic) [27]-[33], [39]. Not only can the method approximate the field and compute the cost function, it can also be used to find the cost function sensitivities (gradients) with respect to the design parameters, needed for the gradient-based optimization. Calculating accurate design sensitivities inexpensively is essential for the efficiency of the gradient-based optimizer. The traditional finite difference approach is computationally expensive, especially for problems that have more than one design parameter. This chapter explains how to obtain the sensitivities over a range of frequencies at a low computational cost, regardless of the number of design parameters. The significance of automatic mesh generation and geometrical parameterization of complex structures in the optimization process are also highlighted. The cost function and sensitivities thus found are used in the optimization of 3D microwave devices with the line search technique outlined in Chapter 2. Results obtained for the test cases are presented in this chapter.

3.1 Parameterization and mesh generation

Design parameterization refers to the process of defining geometrical parameters for a given device that are varied during the optimization. This process is usually left to the designer so that the parameters that have substantial influence on the overall performance of the device can be identified and incorporated into the optimization analysis. The cost function gradient, used by the optimizer to converge to the local minimum, is the rate of change of the cost function with respect to these varying design parameters. The dimension-driven optimization process generates new device geometries at each optimization step in its attempt to improve the targeted device characteristic [11], [34].

In the FE method, the given structure is broken up into numerous non-overlapping elements which constitute the mesh. The degrees of freedom associated with each element (through the nodes, edges and faces, for 3D vector finite elements) determine the FE approximation of the unknown field throughout the structure. During the optimization process, the FE mesh needs to adjust to the varying geometry of the problem, so that the solver can calculate the new field approximation accurately. In order to create the new mesh independently at each optimization step, an automatic mesh generator is required.

However, the FE solver deals with the model in terms of absolute Cartesian co-ordinates of the vertices of the elements, which are rarely useful as geometric parameters. Therefore a 'mapping'[2] between the design parameters of the geometric model and the co-ordinates of the vertices used by the FE solver is necessary to identify the vertices that are directly associated with the design variables. This 'mapping' is critical when calculating the cost function sensitivities with respect to the design variables. At each optimization step, the latest values of the design parameters are incorporated into the geometric model generation process. The model faces are assigned unique labels, so that they can be identified with the associated design parameters. Using the mesh created for the current geometry, the FE solver computes the cost function and its gradients. Computing the cost function gradient requires the sensitivity of every vertex co-ordinate to each design parameter (the vertex sensitivity). Each vertex is identified with the labeled face on which it lies, and consequently, with the design parameter(s) associated

[2] The term 'mapping' is used here to simply denote the correspondence between a design parameter and the mesh vertices

with that face. Once the vertex has been identified, the vertex sensitivity information is computed through a formulation that incorporates the co-ordinate information of the vertex, the dimensional data for the face, as well as the direction and magnitude of the change in the design parameter. This process is repeated for every vertex k along co-ordinate l (x, y, z for the Cartesian system) with respect to the associated design parameter g, yielding the vertex sensitivity ${dr_k^l}/{dg}$. It is done once at the beginning of each optimization step, thus providing the vertex sensitivity information with respect to the *latest* design parameters. These sensitivity values are also used in the mesh perturbation process, as explained next.

The cost function is a continuous function of the design parameter vector $\{g\}$. For small changes in $\{g\}$, the cost function will remain smooth as long as the mesh topology (the node interconnections of the tetrahedral elements) remains the same, even though the geometry (i.e., the node co-ordinates) changes. However, when the changes in the parameters are significant enough, the automatic mesh generator will generate a new topology to better approximate the geometry. This sudden change in the cost function will cause an abrupt jump in the cost function (due to discretization error), making it discontinuous. In 3D problems such changes are frequent, resulting in numerous instances throughout the analysis where the cost function is discontinuous, which will adversely affect the line search efficiency.

To avoid such situations, the optimizer should use the same mesh topology for as long as possible, with only the nodes moved (perturbed) to reflect the changes in the geometry. For example, if the new $\{g\}$ is close enough to the design parameter vector that was used to construct the mesh topology (denoted by $\{g^{Build}\}$), the generator can use the old mesh topology generated for $\{g^{Build}\}$, but with the nodes slightly perturbed to a value equal to the difference between the two. The vertex sensitivity values are used in determining the exact amount by which each node should be moved to approximate the new geometry. On the other hand, if $\{g\}$ is too different from $\{g^{Build}\}$, perturbation could result in elements that are distorted, which in turn could cause the FE analysis to fail, or produce bad solutions. Such situations warrant the generation of a mesh with a new topology using $\{g\}$ to approximate the new geometry properly.

Defining a threshold value for maximum change allowed in $\{g\}$ helps the automatic mesh generator determine the proper course of action. For example, setting a threshold value of $\gamma\%$ would mean that a new mesh needs to be generated if the largest difference between $\{g\}$ and $\{g^{Build}\}$ is greater than $\gamma\%$, i.e., if

$$\max\left\{\frac{\left|g_1^{Build}-g_1\right|}{g_1^{Build}},\frac{\left|g_2^{Build}-g_2\right|}{g_2^{Build}},\ldots,\frac{\left|g_n^{Build}-g_n\right|}{g_n^{Build}}\right\}\times 100>\gamma\% \tag{3.1}$$

where n is the total number of design parameters. The threshold value can be set based on the extent of the change that is deemed tolerable by the designer, taking into account the geometrical properties of the given problem, before generating a new mesh.

For this thesis, the commercial software package ElecNet is used for modeling the problem, as well as for mesh generation [35]. (ElecNet is a 3D finite element system for simulation of electro-quasistatic fields, but just the pre-processing stages are used here.) The optimizer is written in Visual Basic script, and is run within Microsoft Excel [36]. The VB script calls ElecNet to generate the geometrical model, then assign the proper labels to the geometry faces as specified by the designer, and generate the corresponding FE mesh. The mesh data and the geometry-face labeling information are transferred over to the FE solver as data files for further processing and sensitivity calculations.

3.2 Calculating scattering parameters using the FE method

Generally, a microwave component can be viewed as an electrical network with N ports (an N-port device), whose performance is characterized by network parameters (e.g., impedance, admittance or scattering). This representation of the microwave network lends itself to the development of equivalent circuits of arbitrary networks, and is quite useful in the design and analysis of passive components [37].

The scattering parameters provide a representation for a given microwave network that is related to the incident, reflected and transmitted waves. They provide a complete description of the network as seen at its ports, by relating the normalized voltages of waves incident on the ports to the voltages of waves leaving from the ports. The matrix [S] of scattering parameters is defined by the following relation [37]:

$$\{V^-\} = [S]\{V^+\} \tag{3.2}$$

where $\{V^-\}$ is the column vector representing the waves leaving from all ports and $\{V^+\}$ represents the column vector for the incident waves at the ports.

Device performance characteristics such as return loss and insertion loss, which indicate the efficiency of the microwave network, are defined in terms of the scattering parameters. In turn, the scattering parameters can be determined from the field within the device approximated using a numerical analysis technique, for example, the FE method.

3.2.1 S-Parameter calculation from the field

When an N-port microwave device is excited by its dominant mode at port q, with all other ports perfectly matched, the resulting electric field solution in the device can be denoted by $\mathbf{E}_0^{(q)}$. The S-parameters are calculated using this field value, by applying mode orthogonality [38], and using the relation [9]:

$$S_{pq} = V_0^{(p)}(\mathbf{E}_0^{(q)}) - \delta_{pq} \tag{3.3}$$

where δ_{pq} is the Kroenecker delta, and $V_k^{(p)}$ is an operator that extracts the voltage in mode k at port p by projecting against the modal magnetic field $\mathbf{h}_k$:

$$V_k^{(p)}(\mathbf{E}) = -\int_{\mathbf{Port}\ p} \mathbf{E} \times \mathbf{h}_k^{(p)} \cdot \hat{\mathbf{a}}_\mathrm{n}\, dS \tag{3.4}$$

Mode k=0 in (3.4) is the dominant mode, and $\hat{\mathbf{a}}_\mathrm{n}$ is the unit normal vector pointing outward from the port.

Given an accurate field solution for $\mathbf{E}_0^{(q)}$, the scattering parameter matrix for the given microwave device can be calculated.

3.2.2 Finite Element analysis for the field

The field $\mathbf{E}_0^{(q)}$ for an N-port device can be obtained by solving the Helmholtz (curl-curl) equation for the device while satisfying a set of boundary conditions that define the physical characteristics of the problem. The Helmholtz equation can be derived from Maxwell's equations, and is given by [39]:

$$\nabla \times \frac{1}{\mu_r} \nabla \times \mathbf{E} - k_0^2 \varepsilon_r \mathbf{E} = 0 \tag{3.5}$$

where $k_0 = \omega\sqrt{\varepsilon_0 \mu_0}$ is the free-space wavenumber, ε_0, μ_0 are the permittivity and permeability, respectively, of free-space, and ε_r, μ_r are the relative permittivity and permeability, respectively, of the medium.

3.2.2.1 Boundary conditions

Three types of boundary conditions are normally applied in the analysis of microwave devices: Dirichlet, Neumann and Port boundary conditions.

(a) Dirichlet boundary condition: The tangential component of the field over the surface is constrained as:

$$\hat{\mathbf{a}}_{\mathbf{n}} \times \mathbf{E} = \mathbf{E}_{\mathbf{0}} \tag{3.6}$$

where $\hat{\mathbf{a}}_{\mathbf{n}}$ is the unit vector normal to the surface. If $\mathbf{E_0}$ is zero, the surface is a homogeneous Dirichlet boundary condition, which is equivalent to a perfect electric conductor, PEC (short circuit).

(b) Neumann boundary condition: This condition constrains the tangential component of the curl of the field. The homogenous type is most common:

$$\hat{\mathbf{a}}_{\mathbf{n}} \times \nabla \times \mathbf{E} = 0 \tag{3.7}$$

which corresponds to a perfectly magnetic conductor, or PMC.

(c) Port boundary condition: This condition was formulated to model the propagating and evanescent modes present at the ports. Though the main interest is in the dominant propagating mode, the non-propagating modes need to be modeled properly to reduce the error in the computed scattering parameters.

Assuming dominant mode excitation at port q with unit voltage, the boundary condition at port p can be written in terms of the tangential magnetic field as [40]:

$$\mathbf{H}_t = 2\delta_{pq}\mathbf{h}_0^{(p)} - \sum_{l=0}^{\infty} V_l^{(p)}(\mathbf{E}_0^{(q)})\mathbf{h}_l^{(p)} \tag{3.8}$$

where $V_l^{(p)}(\mathbf{E}_0^{(q)})$ is the voltage extracting operator from (3.4), and δ_{pq} is the Kroenecker delta. The summation for l (0->∞) indicates the infinite number of modes (including the evanescent ones). In practice, however, the summation will only include the propagating mode and the first few evanescent modes. This boundary condition acts essentially as an absorbing condition, with the propagating and non-propagating modes 'absorbed' at the port boundary. It is enforced by implementing the *port element* that was described in [40].

Figure 3.1 illustrates the problem domain Ω and the boundary conditions applied to a microwave device. The homogeneous condition for (3.6) (i.e., PEC) is considered, and denoted with Ω_e, and the PMC is denoted by Ω_m. An arbitrary port q, to which the Port boundary condition (3.8) is applied, is also shown.

The wave equation given by (3.5) combined with the boundary conditions describe the field distribution in free space for a given microwave device.

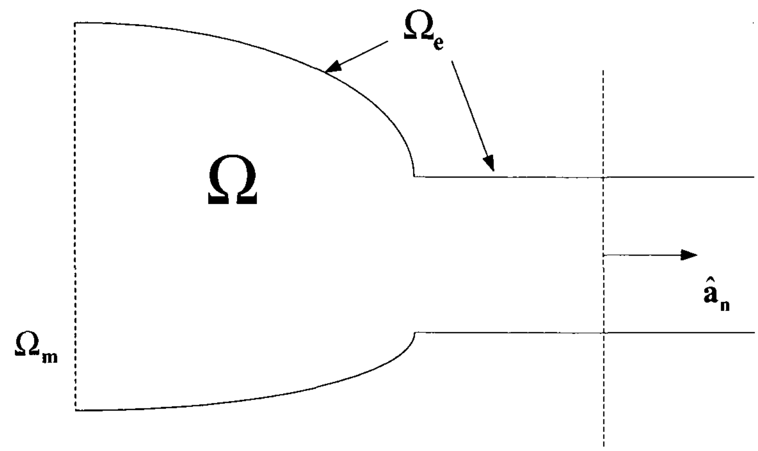

Figure 3.1 Boundary conditions for a microwave device

3.2.2.2 Weak formulation of the problem

To calculate the field distribution by the FE method, the wave equation (3.5) subject to the boundary conditions (3.6)-(3.8) is first replaced by the following equivalent weak formulation:

$$a(\mathbf{E}_0^{(q)}, \mathbf{w}) = b(\mathbf{w}) \tag{3.9}$$

for all weight functions $\mathbf{w}$ satisfying the homogenous version of (3.6), where a is the bilinear form given by:

$$a(\mathbf{E}_0^{(q)}, \mathbf{w}) = \frac{1}{j2k_0\eta_0} \int_\Omega \left(\frac{1}{\mu_r} \nabla \times \mathbf{E}_0^{(q)} \cdot \nabla \times \mathbf{w} - k_0^2 \varepsilon_r \mathbf{E}_0^{(q)} \cdot \mathbf{w} \right) d\Omega + \sum_{p=1}^{N} \sum_{l=0}^{\infty} V_l^{(p)}(\mathbf{E}_0^{(q)}) V_l^{(p)}(\mathbf{w}) \tag{3.10}$$

and b is the linear function given by:

$$b(\mathbf{w}) = V_0^{(q)}(\mathbf{w}) \tag{3.11}$$

where Ω is the volume of the device and η_0 is the intrinsic impedance of free space [40].

Given the definition of the linear functional of (3.11), the S-parameter definition given by (3.3) can be re-written as:

$$S_{pq} = b\left(\mathbf{E}_0^{(q)}\right) - \delta_{pq} \tag{3.12}$$

This definition will become useful when deriving the expressions for the error indicator in Chapter 4.

3.2.2.3 Domain discretization for FE analysis

In order to solve the formulation given by (3.8), the problem domain is divided into non-overlapping sub-domains that constitute the FE mesh. For a typical 3D FE analysis, the mesh could be comprised of tetrahedral, hexahedral or triangular prism elements [39]; in this thesis, tetrahedral elements are used. Within each element, the unknown field is approximated using basis functions, so that the entire domain is represented with a finite number of degrees of freedom. The unknown electric field $\mathbf{E}_0^{(q)}$ can be represented in the discrete form as:

$$\mathbf{E}_0^{(q)} = \{\mathbf{N}\}^T \left\{E_0^{(q)}\right\} \tag{3.13}$$

where $\{\mathbf{N}(x,y,z)\}$ represents the vector basis functions, and $\left\{E_0^{(q)}\right\}$ denotes the corresponding unknown coefficient vector [51].

Substituting (3.13) into (3.9) gives:

$$a\left(\{\mathbf{N}\}^T \left\{E_0^{(q)}\right\}, \mathbf{w}\right) = b(\mathbf{w}) \tag{3.14}$$

Due to symmetry of the bilinear form, (3.14) can be re-arranged as:

$$a\left(\mathbf{w}, \{\mathbf{N}\}^T \left\{E_0^{(q)}\right\}\right) = b(\mathbf{w}) \tag{3.15}$$

Taking the basis function $\{\mathbf{N}(x,y,z)\}$ as the weighting functions, (3.15) is transformed to:

$$a\left(\{\mathbf{N}\},\{\mathbf{N}\}^T\right)\left\{E_0^{(q)}\right\}=b(\{\mathbf{N}\}) \tag{3.16}$$

This system of equations can be written more compactly as:

$$[K]\{E_0^{(q)}\}=\{b\} \tag{3.17}$$

where $[K]$ is a square, symmetric, and sparse global matrix, $\{b\}$ is a known column vector, and $\{E_0^{(q)}\}$ is the global vector of unknowns. The matrices are given as:

$$\begin{aligned} K_{ij} &= a(\mathbf{N}_i,\mathbf{N}_j) \\ b_i &= b(\mathbf{N}_i) \end{aligned} \tag{3.18}$$

This system of equations can be solved for the unknown vectors $\left\{E_0^{(q)}\right\}$ using a direct or an iterative matrix solver to obtain an approximation to the unknown electric field for the given problem domain. For this thesis, the symmetric SOR pre-conditioned conjugate-gradient matrix solver is used [41].

Using (3.13), the scattering parameter expression in (3.12) can be rewritten as follows:

$$\begin{aligned} S_{pq} &= b\left(\{\mathbf{N}\}^T\right)\left\{E_0^{(q)}\right\}-\delta_{pq} \\ &= \{b\}^T\left\{E_0^{(q)}\right\}-\delta_{pq} \end{aligned} \tag{3.19}$$

Substituting the expression for $\{b\}$ in (3.17) into (3.19) gives:

$$S_{pq}=\left\{E_0^{(p)}\right\}^T[K]\left\{E_0^{(q)}\right\}-\delta_{pq} \tag{3.20}$$

This is the discrete form for (3.12), and is used to compute the actual scattering parameter values from the electric field solution.

3.2.3 Computing scattering parameter sensitivities

Calculating the scattering parameters sensitivities with respect to the design parameters cheaply is critical for the efficiency of the gradient-based optimizer. The traditional method of using numerical differentiation, where the sensitivities are calculated by perturbing one design parameter at a time, requires the FE solver to be invoked repeatedly for each perturbed design parameter, resulting in significant computational costs. For an iterative process like the optimization of a 3D problem defined with several design parameters, the cumulative cost becomes prohibitive.

The Adjoint Variable method is an alternate way to obtain the sensitivities cheaply, requiring only one extra solution (in addition to the solution required to compute the unknown field values) to obtain the sensitivities, regardless of the number of design parameters. The extra solution is used to compute the adjoint field, from which the derivative values are calculated. The details of the method can be found in [7]. Specific applications of the adjoint variable method to full-wave electromagnetic analysis of high-frequency structures using FEM are presented in [6], [42]-[43], and the references therein.

Certain modifications to the Adjoint method, specifically when calculating the sensitivities for scattering parameters, were presented in [9]. When the design parameters are internal and do not affect the location and shape of any port of the device, the design sensitivities calculations can be simplified using the following relation:

$$\frac{dS_{pq}}{dg} = -\left\{E_0^{(p)}\right\}^T \frac{d[K]}{dg}\left\{E_0^{(q)}\right\} \tag{3.21}$$

where $[K]$ is the stiffness matrix given by (3.16) and (3.17), and $\{E_0^{(q)}\}$ is the electric field when port q is excited, which is already available from the field calculations using FE.

The derivative of the stiffness matrix $[K]$ with respect to the Cartesian co-ordinates of the vertices of the FE mesh can be obtained as [9]:

$$\frac{d[K]}{dg} = \sum_{k=1}^{3}\sum_{l=0}^{nNodes}\frac{\partial[K]}{dr_k^l}\frac{dr_k^l}{dg} \tag{3.22}$$

where *nNodes* is the number of vertices in the given mesh, and ${dr_k^l}/{dg}$ is the vertex sensitivity value that was explained in Section 3.1. Further details for calculating the sensitivity of the stiffness matrix [K] are given in [43].

With this formulation, the design sensitivities can be calculated directly from the field solutions already needed to compute the full scattering matrix, and no additional system of equations needs solving.

3.3 Gradient based optimization combined with FE

Once the scattering parameters and their sensitivities are calculated using the FE method, they can be used to compute the cost function and its gradient for the optimization analysis. Most designers are interested in the performance of the device over a range of frequencies, rather than at a single frequency point. For the optimization analysis, this requires the cost function and its derivatives be defined over a frequency range. Early attempts to obtain the frequency response over a range involved repeated single frequency analyses at multiple discrete points over the range of interest [44], [45]. However, methods such as Complex Frequency Hopping (CFH) [46], Asymptotic Waveform Evaluation (AWE) [47], and Adaptive Lanzcos-Pade Sweep (ALPS) [48] have proven to be effective in obtaining the response by using moment values of the quantity of interest (e.g. S_{pq}) at a single frequency in a high order expansion over the range. This approach can be used to find the design sensitivities of the frequency response as well, as demonstrated in [49]. This research uses the AWE approach to obtain the frequency response of the scattering parameters and their sensitivities.

3.3.1 Frequency response of S-Parameters

To obtain the frequency response, S_{pq} and its first M moments are calculated at a single frequency point s_0, which can then be used in, for example, a truncated Taylor series approximation about the central frequency s_0, as follows:

$$S_{pq}(s) = \sum_{m=0}^{M} a_m (s - s_0)^m \tag{3.23}$$

where a_m are the moments of S_{pq}, given by:

$$a_m = \frac{1}{m!} \frac{d^m S_{pq}}{ds^m}\bigg|_{s=s_0} \tag{3.24}$$

(3.24) can be evaluated using (3.20) for the m^{th} derivative of $\{E^{(q)}\}$ at s_0, giving the following relation:

$$a_m = \sum_{n=0}^{m} \sum_{l=0}^{m-n} c_{mnl} \left\{E^{(p,n)}\right\}\left[K^{(m-n-l)}\right]\left\{E^{(q,l)}\right\} \tag{3.25}$$

where $\{E^{(q,m)}\}$is the m^{th} derivative of $\{E^{(q)}\}$ with respect to s at s_0 ; $[K^{(v)}]$ is the v^{th} derivative of $[K]$ with respect to s at s_0; and

$$c_{mnl} = \frac{1}{(m-n-l)!\,n!\,l!} \tag{3.26}$$

The derivatives $\{E^{(q,m)}\}$ can be found during the standard frequency analysis for the S-parameters by solving repeatedly a matrix system $[K]\{x\}=\{b\}$, with a different $\{b\}$ each time [46].

For the frequency response sensitivities, the following relation can be obtained from (3.23) and (3.24):

$$\frac{dS_{pq}(s)}{dg} = \sum_{m=0}^{M} \frac{da_m}{dg}(s - s_0)^m \tag{3.27}$$

and

$$\frac{da_m}{dg} = \frac{1}{m!} \frac{d^m}{ds^m} \frac{dS_{pq}}{dg}\bigg|_{s=s_0} \tag{3.28}$$

The coefficients in (3.27) can be computed using (3.21), and expressed as:

$$\frac{da_m}{dg} = -\sum_{n=0}^{m} \sum_{l=0}^{m-n} c_{mnl} \left\{E^{(p,n)}\right\} \frac{d\left[K^{(m-n-l)}\right]}{dg} \left\{E^{(q,l)}\right\} \tag{3.29}$$

3.3.2 Pade approximation for the frequency response values and sensitivities

A better method to obtain the frequency response values than the Taylor series formulation in (3.23) is the Pade approximation, where the response and the sensitivities are approximated using the function:

$$\frac{\sum_{l=0}^{P} b_l (s-s_0)^l}{1+\sum_{l=1}^{P} c_l (s-s_0)^l} \tag{3.30}$$

A detailed procedure to compute the coefficients for the Pade approximation function given by (3.30) from a_m or da_m/dg is given in [50], and is used to obtain the frequency response for this research.

3.3.3 Computing the cost function and its gradient

The cost function, expressed as a function of the design parameters, should fully characterize the device. For this research, the cost function is defined in terms of the return loss for the given passive microwave device. The optimizer tries to improve device performance by finding the design configuration that will reduce the return loss the most.

Define the following function in terms of the frequency, f:

$$c(f) = |S_{11}(f)|^2 \tag{3.31}$$

for some frequency range $f_1 < f < f_2$. The cost function sensitivity with respect to a design parameter g is then given by:

$$\begin{aligned}\frac{dc(f)}{dg} &= \frac{d}{dg}|S_{11}|^2 \\ &= S_{11}\frac{dS_{11}^*}{dg} + S_{11}^*\frac{dS_{11}}{dg} \\ &= S_{11}\frac{dS_{11}^*}{dg} + \left(S_{11}\frac{dS_{11}^*}{dg}\right)^*\end{aligned}$$

$$= 2\,\mathrm{Re}\left(S_{11} \frac{dS_{11}}{dg}^{*} \right) \quad (3.32)$$

The FE analysis gives the complex values S_{11} and dS_{11}/dg at the central analysis frequency f_0. Using the Pade approximation technique, the values for S_{11} and dS_{11}/dg are calculated at each of N_f sample frequency points across the desired range $f_1 < f < f_2$. The number of sample points, N_f, is specified by the designer to approximate frequency response at a given design point. Following this, the cost function and gradient value at each frequency point are obtained using (3.31) and (3.32). A cost function and gradient definition, $C, \nabla C$, characterizing the behavior of the device over the entire frequency range, can be formulated based on these values, and used for the optimization analysis outlined in Chapter 2.

3.3.4 Optimization and FE mesh generation

The outline of the modified steepest descent algorithm was explained in Chapter 2, and is repeated here, with the FE method used for the numerical analysis module in Step 1:

0. Set threshold angle θ_{th}; and initial point in design space
1. *Generate the mesh for the design configuration, and solve the problem using FE.*
2. If the current point is the initial one from Step 0, set the line search direction to $-\nabla C$; go to 5.
3. If the angle between the new $-\nabla C$ and the current line search direction exceeds θ_{th}, go to 4; else go to 5.
4. Start a new line search with search direction set to $-\nabla C$.
5. Find the new design point along the current line; go to 1.

The FE method is part of the CFE module of the line search algorithm. It simply produces the values of $C, \nabla C$ using uniform order elements and the mesh discretization corresponding to the latest design at any stage of the line search. The mesh generation block is introduced as part of the CFE module, taking the latest design parameter values from the optimizer and generating the approximate mesh needed for the FE analysis. The mesh generation is a slightly complex process due to issues that were outlined in Section 3.1, with implications on the efficiency of the optimization.

The modified version of the flow diagram in Figure 2.3 is shown in Figure 3.2. It illustrates the combined FE analysis-optimization scheme, with emphasis on the FE mesh generation process. The design parameter vector $\{g\}$ corresponding to the initial geometry of the pre-optimization analysis is set as the reference value $\{g^{Build}\}$, which in turn, is used to evaluate mesh perturbation conditions (Equation (3.1)) for future design points. Section 2.4 explained how the cubic line search is implemented with step length v, and design parameter vectors $\{g\}$, $\{g^{Start}\}$ and $\{g^{Bound}\}$. If the latest design point $\{g\}$ satisfies the condition in (3.1), a new mesh topology is required to better approximate the new geometry. The optimizer terminates the current line search, and a new one is initiated using the latest design vector $\{g\}$ as the starting point (i.e., the lower bound). This ensures that the same mesh topology is maintained to approximate the geometry during the course of a given line search, maintaining a smooth cost function. In the absence of a re-mesh scenario, the angle criterion is evaluated to determine whether or not to terminate the line search, as explained in Section 2.3.

The mesh generation block is positioned between the optimizer and the FE analysis module (shown as part of the CFE module in Figure 3.2), details of which are shown in Figure 3.3. Note that mesh perturbation is not necessary for the pre-optimization step, as it is the initial point of the analysis, and the design parameter vector is set as the initial reference vector $\{g^{Build}\}$. For the new design point $\{g\}$, if the change in the design parameters (as per (3.1)) is insignificant, the old mesh based on $\{g^{Build}\}$ is used for the geometry and perturbation is applied with the help of the vertex sensitivity values computed for the latest design parameter values. The $\{g^{Build}\}$ values remain unchanged, as it still provides a valid mesh topology to discretize the geometry. On the other hand, a considerable change in the design parameter values will lead to an update of $\{g^{Build}\}$. The mesh generation block will alert the optimizer of the re-mesh scenario through a flag, prompting it to terminate the line search after calculating the latest $C, \nabla C$.

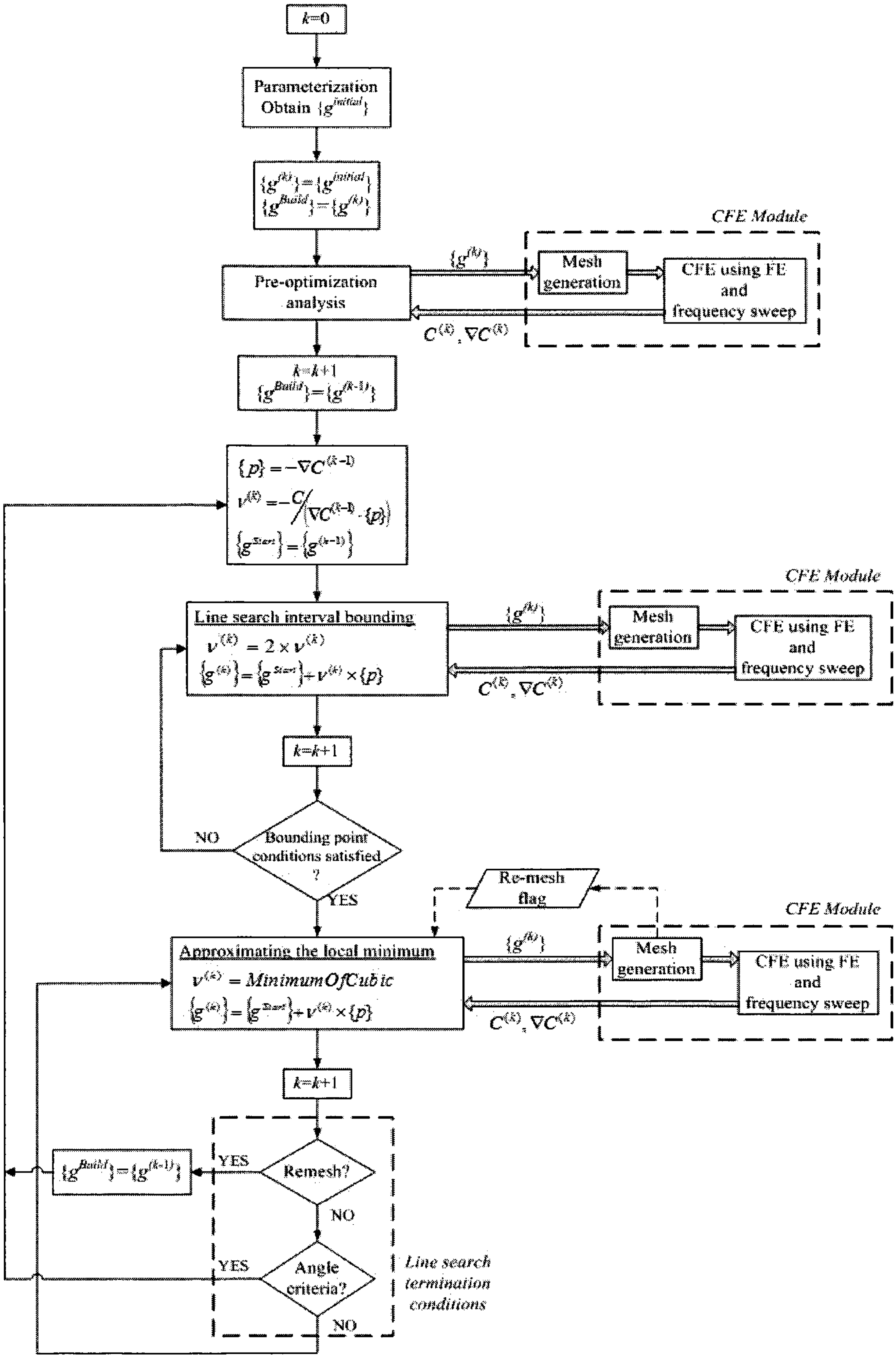

Figure 3.2 Flow diagram for the combined optimiztion-FE analysis

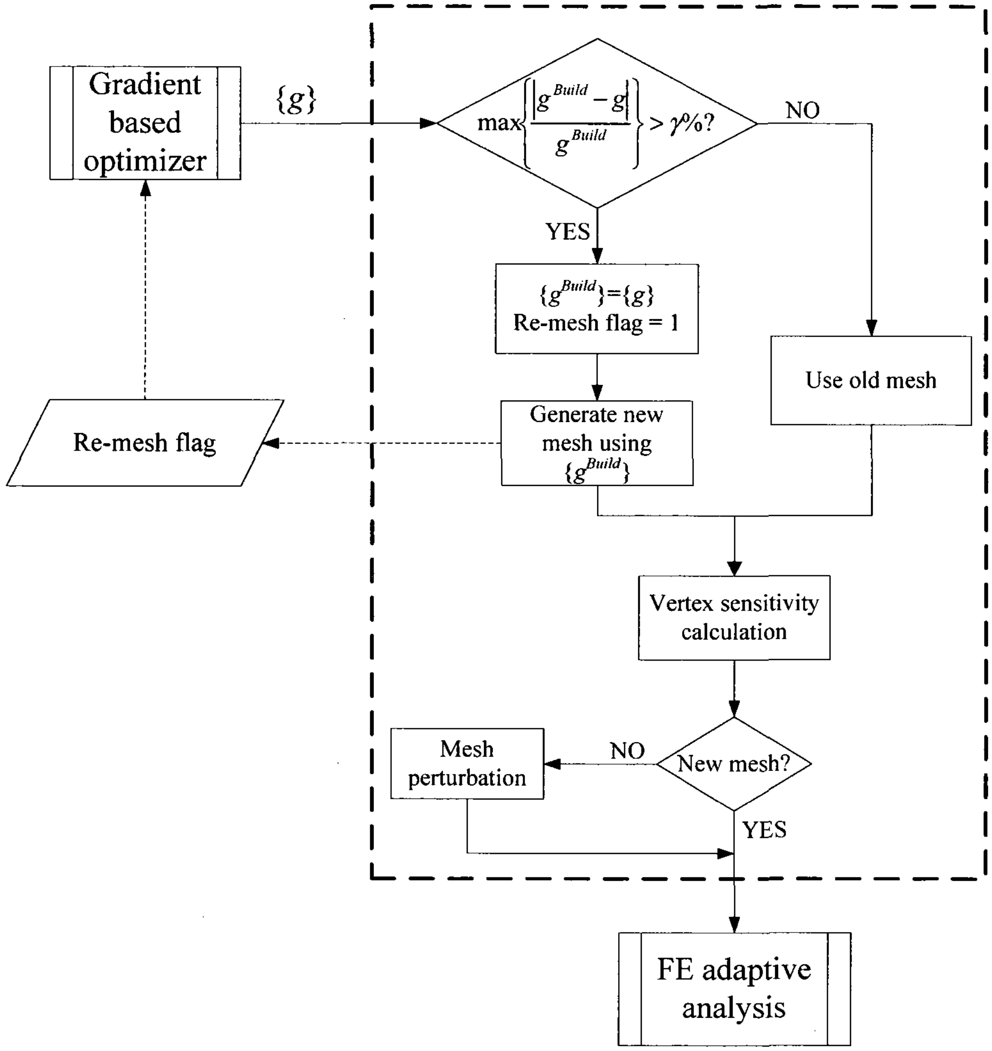

Figure 3.3 Mesh generation block for cubic point analysis

The vertex sensitivities are calculated for every step of the analysis, in order to update the values with respect to the latest geometry, and consequently compute the cost function gradient values accurately. They are also used in the mesh generation process, specifically for mesh perturbation, where they are used to adjust the old topology to the latest changes in the design parameters.

3.4 Test Cases and Results

The combined optimization-FE analysis scheme illustrated by Figure 3.2 was tested on three two-port waveguide components. Each problem was excited by the dominant TE_{10} mode, and the operating frequency range was selected so that only the dominant mode is allowed to propagate. The Dirichlet boundary condition given by (3.6) was applied to all the waveguide walls, while the Port boundary condition of (3.8) was applied to the ports. Uniform high order elements based on the hierarchal basis functions described in [51] were used for the FE analysis to obtain the field solutions. Using the Pade approximation, the values of c given by (3.31), were obtained at each of N_f=1024 points across the range $f_1 < f < f_2$, from which C and ∇C were obtained. For the test cases presented here, the following cost function and gradient definition was used:

$$C = \max_{f_1 < f < f_2} c(f) \tag{3.33}$$

The value of f where the cost function C is defined (i.e., the maximum value of c in the entire range) is denoted as $f_{\max}$. The corresponding gradients are calculated as:

$$\frac{dC}{dg} = \frac{dc(f_{\max})}{dg} \tag{3.34}$$

These values are then used by the cubic line search analysis to update the value of $\{g\}$. The re-mesh condition of (3.1) was implemented using γ = 10%. The results were reported in [52].

3.4.1 E-Plane Miter bend

A mitered, right-angled E-plane bend shown in Figure 3.4 is the first test case for the combined optimization scheme. The single design parameter g controls the mitering of the bend, and is allowed to vary from g/b =0 to g/b=1 (no mitering) to find the value of g that will give the optimal device performance. The guide cross-section is a x b, with a=2b, b=10mm. The operating frequency range was selected as 8.25 GHz < f < 13.5 GHz.

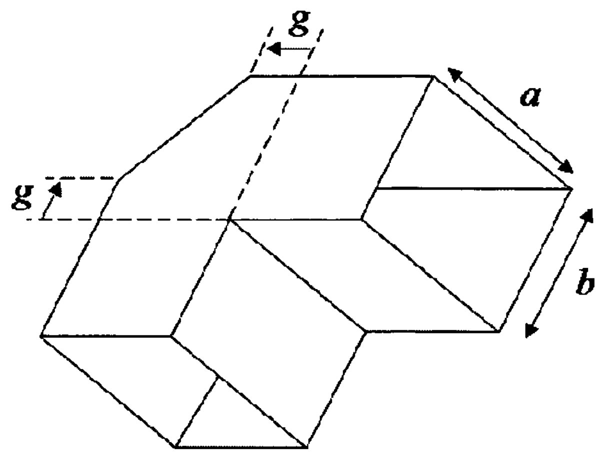

Figure 3.4 The E-plane miter bend with a single design parameter. Starting value for g=1 mm; Central analysis frequency= 10.875 GHz; No. of tetrahedra = 264; [9]

The optimization result is shown graphically in Figure 3.5. The analysis is able to achieve an improvement of approximately 15 dB over six line search steps. Since the problem has only one design parameter, all of the steps for the optimization are the result of the cubic line search along a single line, with each cubic point representing a new geometry in design space with the updated value of $\{g\}$.

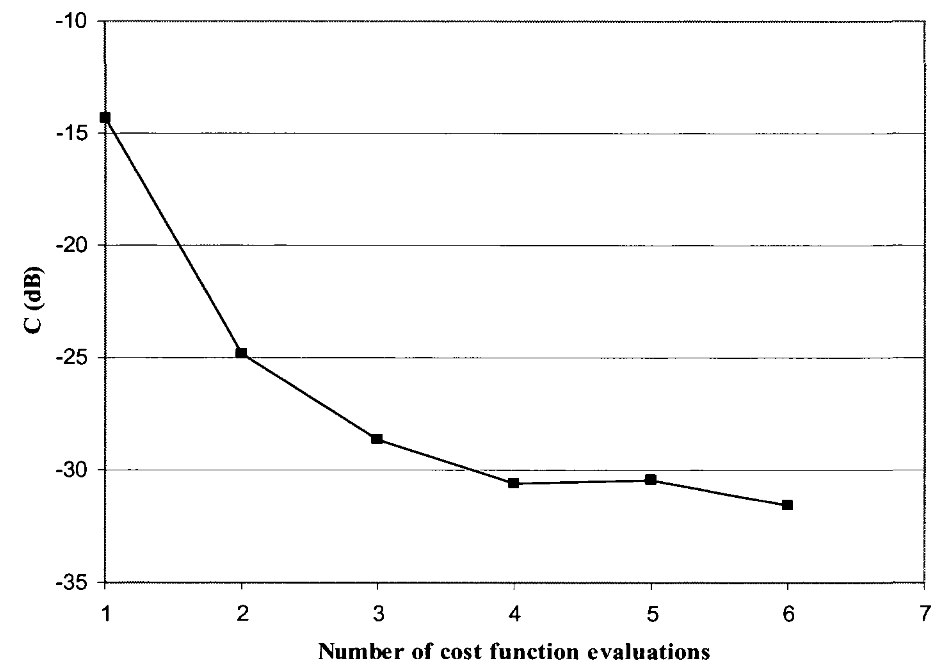

Figure 3.5 Cost function value vs. number of cost function evaluations for the E-Plane bend

The optimizer manages to reduce the cost function value considerably within the first three steps, after which the value appears to have more or less settled down. Table 3.1 shows the variation in C for each new design point at various stages of the optimization. Whereas the worst value in the frequency band for the initial design point is -14.3 dB, the final design point gives a worst value of -31.5 dB. The value for g also appears to have converged to a value close to 2.1mm by the end of the sixth optimization step. Figure 3.6 shows the changes in the frequency response at different points during the course of the optimization. The significant change in the response curves from the initial step to the third step shows the quick improvement attained by the optimization process.

Table 3.1 Design parameter (in mm) and cost function value (in dB) at each optimzation step for the E-Plane over a frequency band and uniform high order elements

Step	g (mm)	C (dB)
1	1	-14.30
2	2.449	-24.85
3	2.237	-28.64
4	2.149	-30.60
5	2.094	-30.43
6	2.121	-31.54

3.4.2 U-bend

The second test case was the symmetric E-plane U-bend, described in detail in [53]. Four geometric parameters were selected as design parameters, as shown in Figure 3.7. Three of the geometric parameters control the mitering, while the fourth (g_4) controls the distance between the arms of the bend. The operating frequency was selected to be 10-15 GHz, within which the optimizer tried to reduce the reflection at either port, using the cost function and gradient definitions given by (3.33)-(3.34).

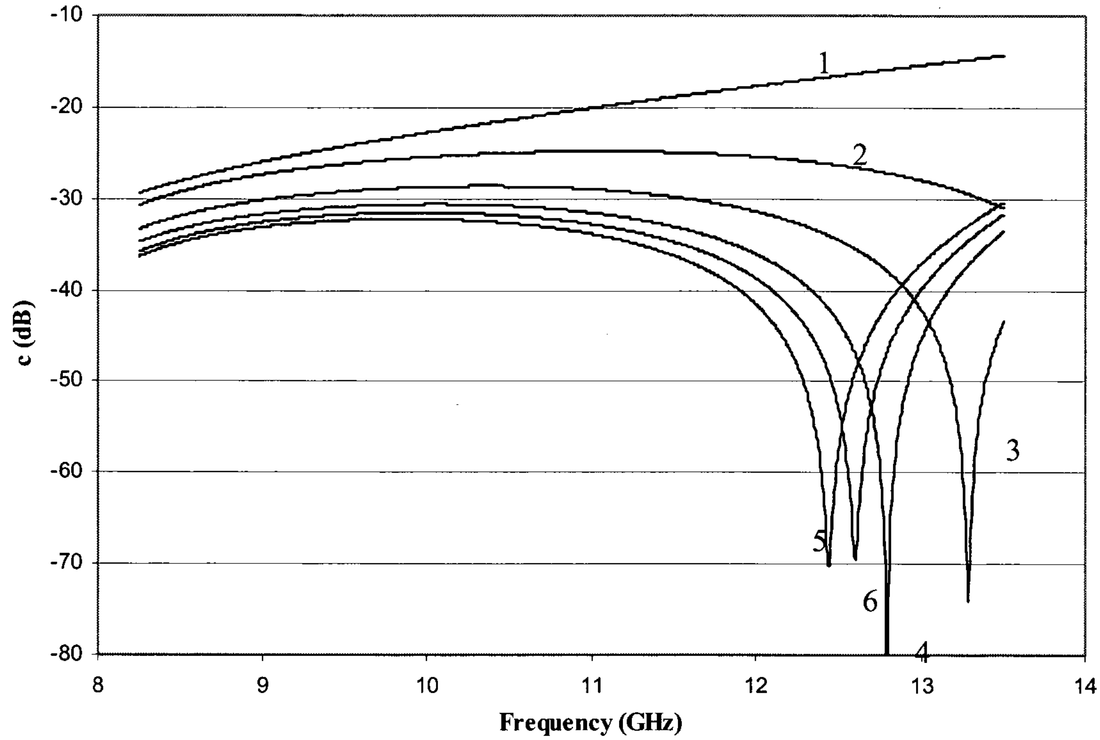

Figure 3.6 The function c for the E-plane miter bend for six successive optimization steps

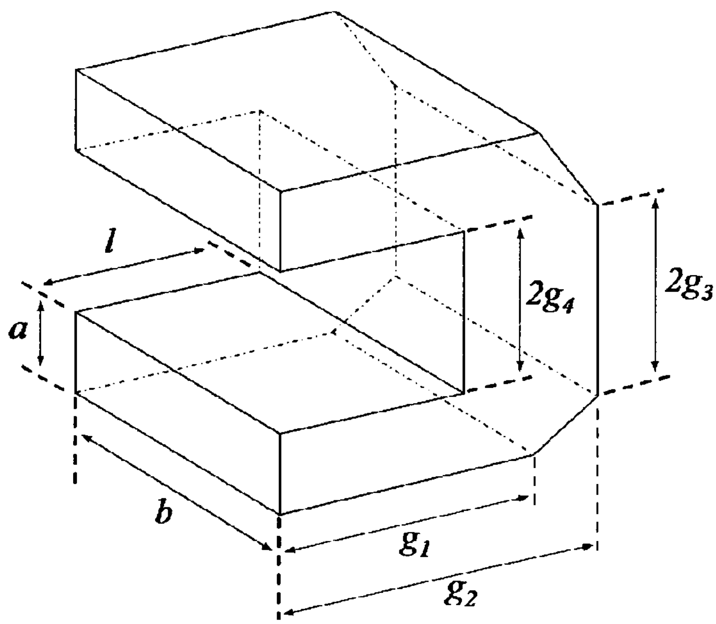

Figure 3.7 An E-plane U-bend with four design parameters. a=5 mm, b=19.05mm, l=10mm; $\{g\}$ = {11.0, 15.0, 3.7, 3.0}mm; Central analysis frequency=12.5 GHz; No. of tetrahedra = 600; [53]

The optimization is now in 4-dimensional space, and multiple line searches were required to reach the local optimum. The cost function reduction over the various line search stages is shown in Figure 3.8. It can be seen that the cost function value gets reduced by approximately 8dB in the first line search with eight CFEs, after which it reduces by about 4dB over 14 CFEs. The optimizer does not seem to be able to obtain any more improvements beyond this point. The spikes in the cost function value at the beginning of each line search in Figure 3.8 represent the line search bounding points.

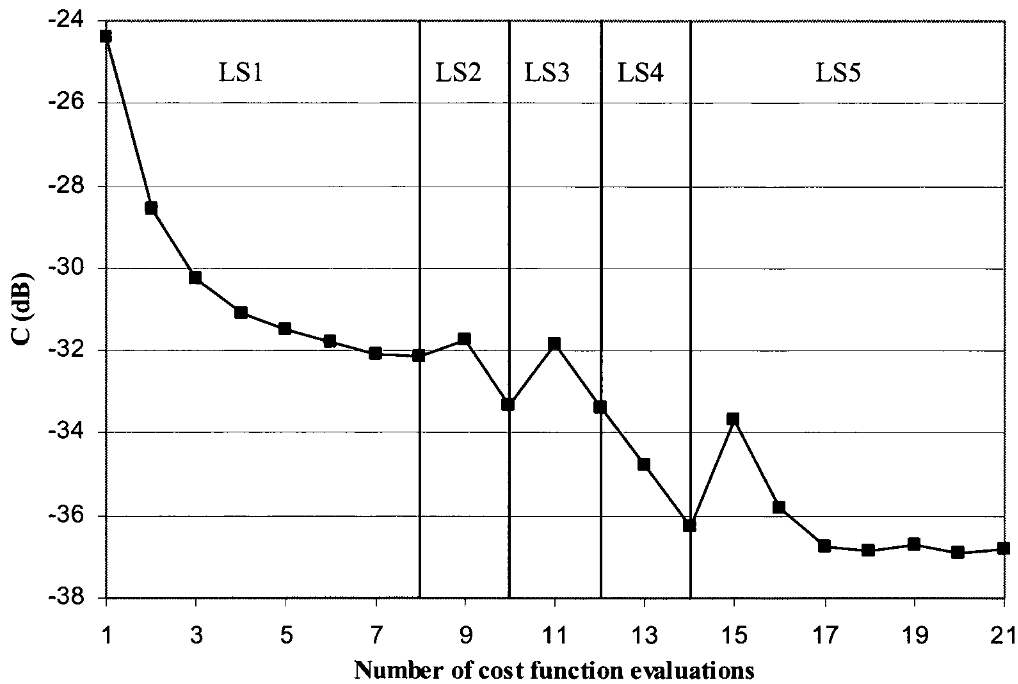

Figure 3.8 Reduction in cost function value for the U-bend problem over 5 cubic line search

Table 3.2 gives a comparison of the initial design parameter values with those obtained at the end of line search step #1 and the final (line search #5) values.

Table 3.2 Design parameter values (in mm) at various steps of the optimization analysis for U-bend

	Initial	Line Search 1	Final
g_1	10.83	10.96	11.01
g_2	15.09	15.03	15.05
g_3	3.75	3.66	3.67
g_4	3.16	3.08	3.04

The change in the design parameter values is minimal, yet significant enough to attain a cost function reduction of close to 12 dB. It can be assumed that the return loss characteristic of the design is highly sensitive to the changes in all the design parameters. The plot in Figure 3.9 shows frequency response of the device over the various optimization steps. For clarity, only the initial curve, the curve at the end of the first line search and the curve from the final line search is shown.

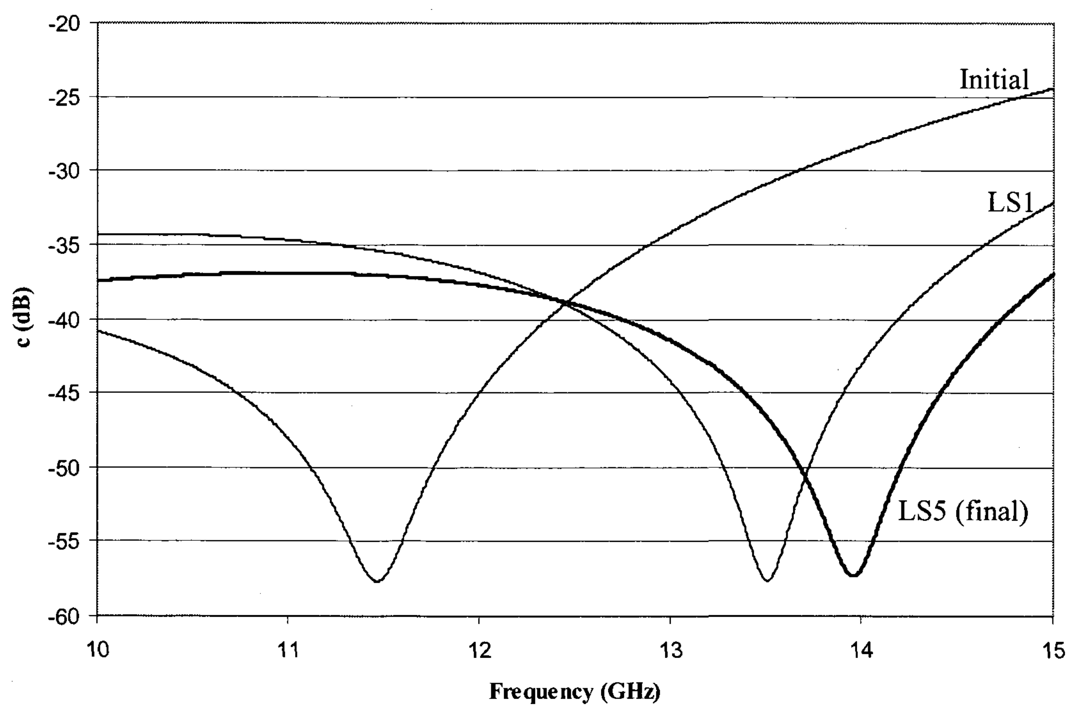

Figure 3.9 Frequency response curves of the E-plane U-bend for the initial line search, and at the end of line searches #1 and #5 (final)

From this plot, the most dramatic change in the frequency response occurs after the first line search, where the optimizer has brought down the worst cost function value point in the band considerably. The final curve shows that the optimizer has achieved a balance between the two end points of the frequency range, from where presumably the cost function values (i.e., the worst points in the band as per the definition in (3.33)) were obtained. The optimizer may have 'tuned' the design parameters to position the resonance dip so as to balance the two end points.

3.4.3 Waveguide impedance transformer

The final test case problem is a waveguide transformer that connects a large cross-section rectangular waveguide to a smaller one, as shown in Figure 3.10. Ideally, there should be no reflections at either port. The performance of the device is governed by the geometry of the middle section, defined by three design parameters $\{a,b,l\}$. The objective is to minimize the reflection at port 1 (larger port: $a_1 \mathrm{x} b_1$), operating over the range 9.75GHz to 10.25 GHz. The initial length of the middle section was selected to be a quarter wavelength at the center frequency (f_0=10GHz), while the other two parameters were chosen to be the average of the corresponding end section values ((a_1, b_1) and (a_2, b_2)).

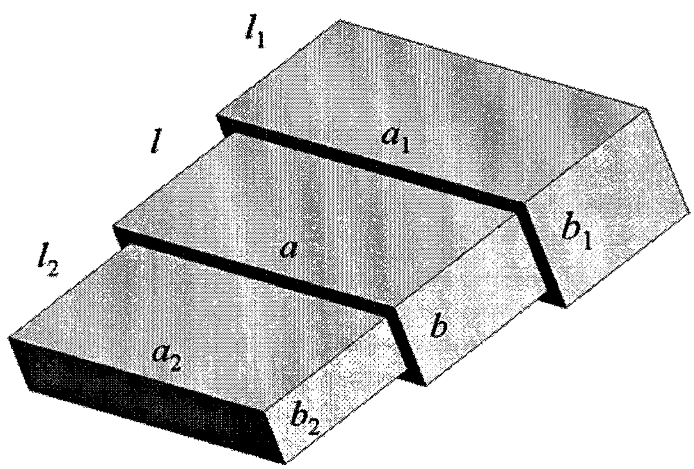

Figure 3.10 Waveguide transformer. a_1**=24mm,** b_1**=8mm,** a_2**=20mm,** b_2**=4mm,** l_1**=10mm,** l_2**=10mm,** a **= 21.70 mm,** b **= 5.68 mm,** l **= 10.30 mm. Design parameter vector** $\{g\} = \{a, b, l\}^T$**; Analysis frequency=10GHz; No. of tetrahedral elements = 1286; [9]**

The initial design configuration provides a decent performance with a return loss of -25dB. However, the optimizer is able to fine tune the design and obtain a further reduction of 7dB. This fine tuning process over six cubic line search steps is given in Figure 3.11, where the cost function is shown to reduce at a steady pace. Just as was the case for the U-Bend problem, the sudden increase in the cost function values at the beginning of each line search is due to the line search interval bounding point analysis.

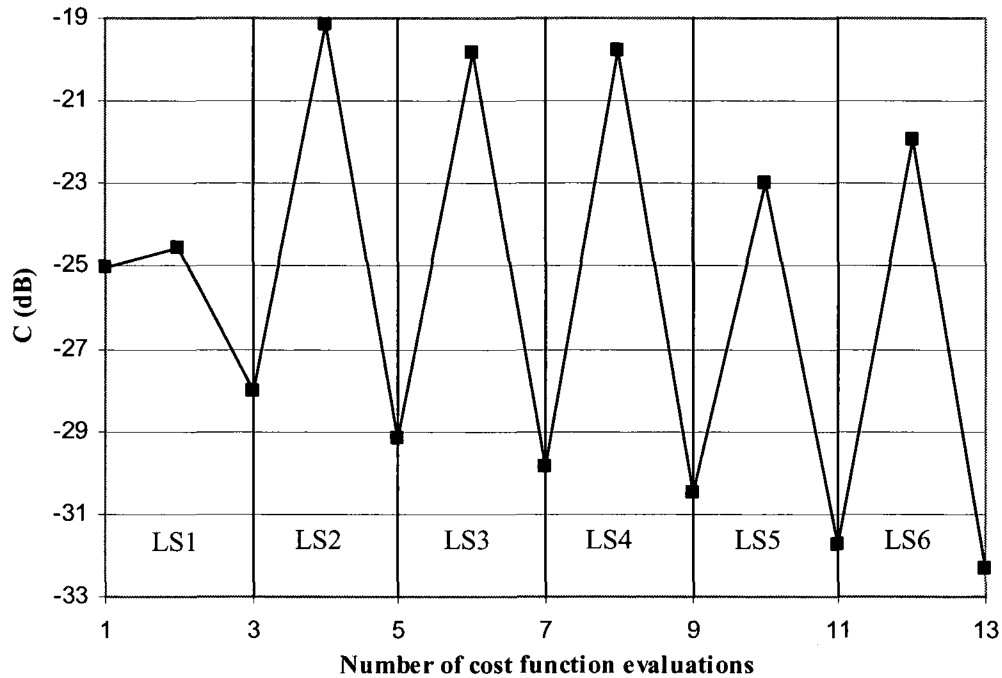

Figure 3.11 Reduction in the cost function for the transformer example

Table 3.3 shows the initial and final values of the design parameters for the transformer. The final values do not show much change from the initial ones, and yet fine-tuning by the optimizer manages to improve the device performance by 7dB over 6 line searches.

Table 3.3 Initial and final design parameter values (in mm) for the waveguide transformer

	Initial	Final
g_1	22.0	21.66
g_2	6.0	5.71
g_3	10.44	10.3

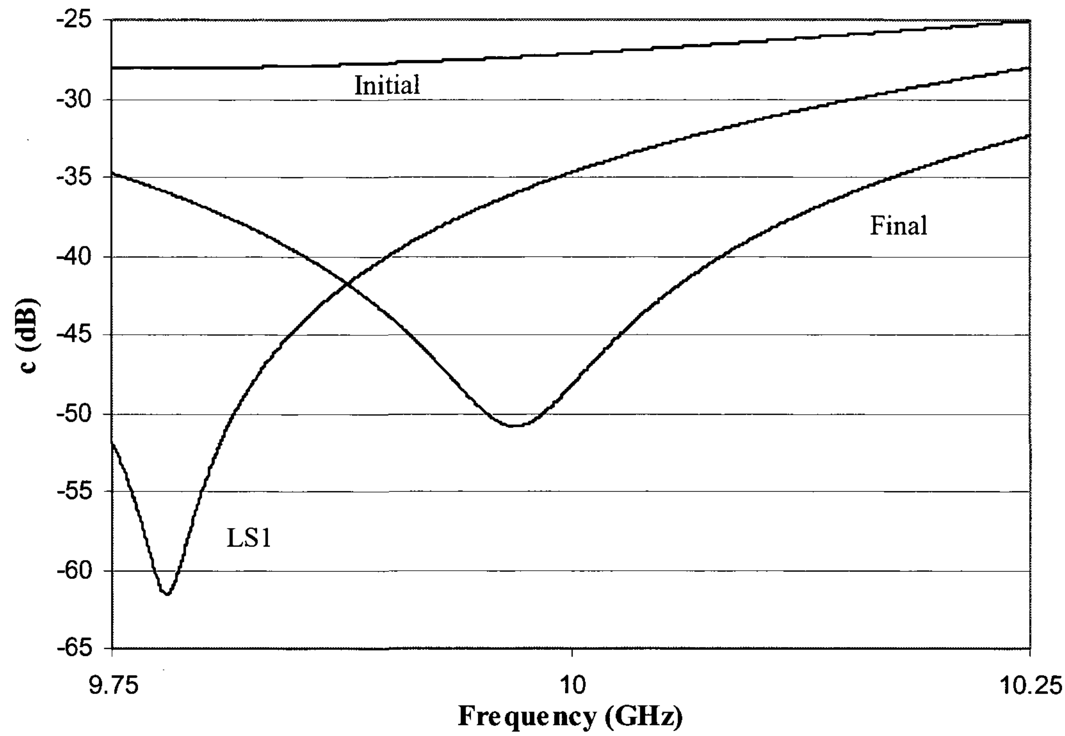

Figure 3.12 Frequency response curves for the waveguide transformer at the end of the initial and final line searches, as well as line search #1.

Figure 3.12 shows the frequency response curves of the transformer for the initial and final steps of the optimization analysis. The optimizer tries to fine-tune the device performance by positioning the resonance frequency point within the range to attain a balance between the end points of the range, where, based on the plots, the worst cost function values occur. By doing so, the optimizer also manages to improve the return-loss performance of the device throughout the operating frequency range, as the final design curve (in contrast to the initial curve) shows.

3.5 Discussion

Results in this chapter validate the use of a cost function and gradient defined over a frequency range in the optimization of 3D microwave problems. The Pade technique has been shown to be effective in obtaining the cost function and gradient values at any sample point in the desired band. Computing the sensitivities using the FE method inexpensively at the central frequency point is an added attraction when applying the method to realistic 3D microwave problems. Special attention is also given to the mesh generation process, where modifications helped in maintaining cost function continuity along a line search. A special line search termination criterion based on the smoothness of the mesh generated is also used for the same purpose.

The results are, however, based on the FE analysis using uniform high order tetrahedral elements. Although the results obtained are highly accurate, the computational costs are significant, and may not be desirable from a 3D design perspective, as the designer is always looking for ways to cut cost. Using an adaptive FE analysis scheme might result in relatively accurate results, but at much less computational cost. How this can be achieved using adaptive FE is demonstrated in the next chapter.

4. *p*-Adaptive FE analysis using hierarchal elements

Conducting 3D FE analysis using uniformly high-order elements produces results that are highly accurate, but computationally expensive, and usually unnecessary in the early stages of the optimization. Devising an FE scheme capable of assigning the degrees of freedom based on the error requirements at each optimization step and attaining the desired solution accuracy with the least number of degrees of freedom possible is clearly desirable.

The accuracy of an FE solution is directly dependent on the number of free parameters used to represent the unknown field, and how effectively these parameters are distributed over the problem space. Adaptive FE analysis is a practical way of achieving this cost-efficient distribution, whereby an iterative process assigns the required number of DOFs throughout the mesh so as to obtain solutions with the desired accuracy.

A critical part of the adaptive process is identifying the high error areas (due to inadequate discretization) within the finite element mesh, so that DOFs can be assigned to reduce the error in those regions and improve overall solution accuracy. There are a variety of error indication methods in the literature which are used in conjunction with adaptive schemes [54] -[57]. Adaptive techniques also require accurate error estimates of the global quantity in order to determine whether or not the iterative process has attained the desired solution accuracy [58]. The error indicator and the estimator work in tandem to help the adaptive analysis produce accurate field approximations in a computationally efficient manner.

This chapter presents a new error indicator and a new error estimator for the *p*-adaptive FE analysis of scattering parameters for 3D microwave devices, which, in turn, forms the engine that drives the combined optimization-FE analysis scheme. The hierarchal elements developed by Webb [51] that are used for the *p*-adaptive analysis in this research are also presented, briefly, in this chapter.

4.1 FE adaptive schemes

The use of adaptive techniques is widespread in FE when trying to obtain cost efficient field approximations [59]. There are four basic adaptive schemes:

(a) *h*-type adaption: In this model the refinement of the FE discretization involves increasing the number of elements in the mesh [60]. This method is particularly useful when dealing with mathematical singularities due to sharp material edges and corners, where a larger number of smaller elements are required close to the singularities, but fewer, larger elements will suffice further away [61].

(b) *p*-type adaption: The *p*-adaptive model updates the discretization by increasing the polynomial order of the basis functions while holding the element sizes constant. This avoids the cost of re-meshing associated with the *h*-adaption model, resulting in computational cost savings, especially for 3D problems [62], [63]. The *p*-type scheme is also known to have specific advantages when applied to high frequency problems, where it is able to capture the considerable wave-like variations of fields in homogeneous regions away from singularities in microwave and optical devices effectively, without the help of a highly refined mesh [54], [64].

(c) *hp*-type adaption: The combined *h*- and *p*-type adaption method tries to exploit the advantages of both models, reducing the element size in certain regions, while increasing the order of the elements in other parts of the mesh [65]. However, the implementation and control of a hybrid *hp*-type scheme is complex due to the subtleties related to the coupling of the *h*- and *p*-type adaptive methods. There has been extensive research in finding systematic ways for properly implementing this powerful hybrid technique [66], [67].

(d) *r*-type adaption: The mesh refinement for this method is accomplished without adding any new DOFs to the FE mesh. The element vertices are repositioned in order to improve the accuracy of the computed solution [68]. This scheme is particularly useful for time-dependent problems.

For all adaptive techniques, the underlying procedure is the same, although different techniques are used for the mesh refinement stage of the adaptive loop. A rough outline for any adaptive analysis can be given as:

1. Generate initial mesh.
2. Solve the problem using the FE method.
3. Check for solution accuracy; if adequate, stop adaption.

Otherwise

4. Identify regions of high solution error.
5. Refine discretization by distributing additional degrees of freedom where necessary.
6. Go to 2.

Two important components for this scheme are error indication and error estimation. A clear distinction on the role of the indicator and the estimator in adaptive FE analysis was made in [69], where the indicator was defined as providing an assessment, or ranking, of the error in a given element relative to other elements in the mesh (step 4), while the estimator provides an accurate estimate of the error in a quantity of interest for the overall mesh (step 3).

For this thesis, the *p*-adaptive method is used for the FE analysis. The hierarchal elements developed in [51] for 3D microwave problems were utilized for the implementation of the adaptive method. After a brief description of the hierarchal elements, the rest of the chapter deals with the theory and implementation details of the novel error indicator and error estimator schemes formulated specifically for the *p*-adaptive analysis of the scattering parameters from 3D microwave problems.

4.2 Hierarchal tetrahedral elements

High order polynomial basis functions for general electromagnetics are well established for 3D FE analysis [39]. Standard Lagrangian elements require that same-order basis functions be used over the entire mesh to ensure continuity of the FE solution [70]. This requires turning up the order of all the elements in the mesh simultaneously for the *p*-

adaptive scheme. However, this can be computationally expensive, as a uniform element order increase introduces a very large number of DOFs, especially for 3D problems. A more efficient way would be to increase the order of the elements selectively in areas with higher solution error, resulting in a mixed order mesh. Hierarchal elements can accommodate this, while maintaining the continuity of the computed solution.

Hierarchal elements have been used extensively in FE since they were first proposed by Zienkiewiscz et al. [71], and have found wide application in both low frequency and high frequency electromagnetics [54], [72]. In hierarchal elements, the basis functions of an element are a subset of the basis functions of any element of higher order, so that elements that share a common edge or face end up having a common set of basis functions on that edge or face. The field solution continuity is ensured by locally matching the coefficients of the basis functions at edges and faces of neighboring elements [73].

For 3D electromagnetic wave problems, tetrahedral tangential vector finite elements (also known as edge elements) impose continuity of the tangential component of the field between adjacent elements, while leaving the normal component unconstrained. This relaxation of the normal component continuity has a number of advantages: it plays an important part of addressing the problem of spurious solutions (numerical solutions that are an approximation of a non-physical solution) in FE analysis, it helps in enforcing the correct boundary and interface conditions, and it improves the field modeling around singularities [74], [75].

In order to be able to apply *p*-adaptive FE analysis, edge elements of several orders are needed. The work in [51] provides general expressions from which any order of element can be obtained in a systematic way. The basis functions that define the elements are partially orthogonalized to increase their degree of linear independence and to improve the conditioning of the system matrix which, in turn, helps improve the efficiency of the matrix solver. The elements have two polynomial orders, one for the gradient subspace, G_m (comprised of curl-free functions), and one for the rotational subspace, R_n. Each element has basis functions that span a space G_m+R_n, where m is the order for the gradient subspace and n is the order for the rotational subspace. Following this notation, the element order can be denoted as (m,n). Basis functions for element orders (0,1), (1,2) and (2,3) have been generated, with order (0,1) being the Whitney edge element with one degree of freedom per edge [51]. For brevity, the (0,1), (1,2) and (2,3) orders will be denoted as Orders 1,2 and 3, respectively, from this point onwards in this thesis.

4.3 Error indication

The objective of any FE adaptive scheme is to obtain a solution of desired accuracy with the least computational cost. Error indicators assist the adaptive method in the selective refinement of the FE mesh by targeting the high error areas to improve the accuracy of the solution in a more computationally efficient manner. Several error indicators have been presented for the FE analysis of electromagnetic problems in the literature [76]-[81]. Although the indicators described in these papers were proven to be effective in guiding the adaptive analysis, they were designed to improve the accuracy of the electric fields or magnetic fields as quickly as possible. From a design perspective, the quantities of interest do not necessarily have to be the fields themselves, but rather quantities that better characterize the behavior of the microwave component, for example, scattering parameters. In addition, since a cost function defined in terms of the scattering parameters of a microwave device is better suited for design optimization purposes (as was done in Chapter 3), an indicator that targets these parameters would be conducive to the overall efficiency of the optimization.

Targeted, or goal-oriented error indicators, where the elements are chosen for refinement based on their contribution to a specified global quantity, have been shown to be more effective than general, field-based error indicators in many cases. The targeted error indictors are sensitive to the desired global quantity, and focus on areas in the FE mesh that will improve the accuracy of that quantity. In [82], a general framework for targeted error indicators was outlined, and was applied to two-dimensional (2D) low frequency electromagnetic problems, where it was shown to be more effective than the general indicators. More recently, work done by Sun et al [83] saw the application of a goal-oriented error indicator targeting the scattering parameters in a 3D h-adaptive analysis scheme. The approach taken was to first obtain an estimate of the field error in the element, and then use it to estimate the element's contribution to the error in S_{11}.

Using the hierarchal elements and p-adaption, a more direct approach to error indication for scattering parameters can be pursued. An error indicator technique outlined in [84], denoted as Scheme 1, involves setting the coefficients of the highest order basis functions to zero at a given adaptive step, and then re-calculating the contribution of the element to S_{11}, with the indicator being the resulting change in S_{11}. However, an improved scheme, denoted as Scheme 2, where the indicator *added* basis functions of the next highest order (instead of subtracting) to find the element's contribution to the error in S_{11}, was also outlined and shown to be more effective than Scheme 1 [82]. This thesis uses the idea from Scheme 2 to formulate a 'forward-looking' error indicator that targets the scattering parameters of 3D microwave problems.

For the finite element problem built using the hierarchal elements from [51], two error indicators become necessary: one representing the gradient part of the field (ε_{grad}), and the other for the rotational component (ε_{rot}). Thus, for each element at each adaptive step the Scheme 2 error indicator is applied twice, with the total error indicator for that element represented by the sum of the two parts:

$$\varepsilon = \varepsilon_{grad} + \varepsilon_{rot} \tag{4.1}$$

Using the indicator values, all the tetrahedral elements in the mesh are ranked, with the high error elements being at the top of the list. The top α% of the elements are then chosen for refinement [83], where α is a pre-determined value.

4.3.1 A targeted error indicator for scattering parameters

The electric field, $\mathbf{E}^{(q)}$, that results when a wave is incident at port q when all other ports are matched, can be computed by solving the following discrete problem:

$$a(\mathbf{E}^{(q)}, \mathbf{w}) = b^{(q)}(\mathbf{w}) \tag{4.2}$$

for all the weight functions $\mathbf{w}$ on the same finite element mesh, and where a and b are the bilinear and linear forms, respectively, explained in Section 3.2.2.2. The scattering parameters, which are a function of the field, can be expressed as (similar to (3.12)):

$$S_{pq} = b^{(p)}(\mathbf{E}^{(q)}) - \delta_{pq} \tag{4.3}$$

An error indicator ΔS_{pq} for an element can be found by taking the difference:

$$\Delta S_{pq} = \left| b^{(p)}(\mathbf{E}_e^{(q)}) - b^{(p)}(\mathbf{E}^{(q)}) \right| \tag{4.4}$$

where $\mathbf{E}_e^{(q)}$ is the new electric field obtained by increasing the element order and re-solving it, which also happens to be a linear combination of the basis functions. Since the elements are hierarchal, it can be split into two parts: one part comprised of the original basis functions, and the other due to the newly added ones, and can be represented as follows:

$$\mathbf{E}_e^{(q)} = \widetilde{\mathbf{E}}^{(q)} + \Delta\mathbf{E}^{(q)} \tag{4.5}$$

where $\widetilde{\mathbf{E}}^{(q)}$ is the electric field due to the original basis functions, and $\Delta\mathbf{E}^{(q)}$ is the part arising from the additional DOFs.

Substituting this back into (4.4), the following expression for the error indicator is obtained:

$$\Delta S_{pq} = \left| b^{(p)}(\Delta\mathbf{E}^{(q)}) + b^{(p)}(\widetilde{\mathbf{E}}^{(q)} - \mathbf{E}^{(q)}) \right| \tag{4.6}$$

Although an approximate expression for the error indicator in terms of $\Delta\mathbf{E}^{(q)}$ can be obtained by discarding the second term in (4.6), it is possible to come up with a form for the second term that will be *exact* in terms of $\Delta\mathbf{E}^{(q)}$, by introducing the original field $\mathbf{E}^{(p)}$ as follows:

$$\begin{aligned}
b^{(p)}(\widetilde{\mathbf{E}}^{(q)} - \mathbf{E}^{(q)}) &= a(\mathbf{E}^{(p)}, \widetilde{\mathbf{E}}^{(q)} - \mathbf{E}^{(q)}) \\
&= a(\mathbf{E}^{(p)}, \mathbf{E}_e^{(q)} - \mathbf{E}^{(q)} - \Delta\mathbf{E}^{(q)}) \\
&= -a(\mathbf{E}^{(p)}, \Delta\mathbf{E}^{(q)}) + a(\mathbf{E}_e^{(q)}, \mathbf{E}^{(p)}) - a(\mathbf{E}^{(q)}, \mathbf{E}^{(p)}) \\
&= -a(\mathbf{E}^{(p)}, \Delta\mathbf{E}^{(q)}) + b^{(q)}(\mathbf{E}^{(p)}) - b^{(q)}(\mathbf{E}^{(p)}) \\
&= -a(\mathbf{E}^{(p)}, \Delta\mathbf{E}^{(q)})
\end{aligned} \tag{4.7}$$

The expression for (4.6) subsequently becomes:

$$\Delta S_{pq} = \left| b^{(p)}(\Delta \mathbf{E}^{(q)}) - a(\mathbf{E}^{(p)}, \Delta \mathbf{E}^{(q)}) \right| \tag{4.8}$$

A new linear form that characterizes this change in the field can be defined as:

$$c^{(p)}(\Delta \mathbf{E}^{(q)}) = b^{(p)}(\Delta \mathbf{E}^{(q)}) - a(\mathbf{E}^{(p)}, \Delta \mathbf{E}^{(q)}) \tag{4.9}$$

Then, the error indicator expression becomes:

$$\Delta S_{pq} = \left| c^{(p)}(\Delta \mathbf{E}^{(q)}) \right| \tag{4.10}$$

The advantage of the expression in (4.8) is that it is still exact, and yet depends only on the degrees of freedom added when the element order is increased, not on $\widetilde{\mathbf{E}}^{(q)}$.

4.3.1.1 Approximating the error indicator

The field $\mathbf{E}_e^{(q)}$ satisfies the equation:

$$a(\mathbf{E}_e^{(q)}, \mathbf{w}_e) = b^{(q)}(\mathbf{w}_e) \tag{4.11}$$

for all weight functions $\mathbf{w}_e$ on the mesh that has the order of element e increased. If the expression in (4.8) is used as the error indicator, it becomes necessary to find $\mathbf{E}_e^{(q)}$ to obtain $\Delta \mathbf{E}^{(q)}$, and solving (4.11) to obtain the exact field is much too expensive. Hence, an approximation to $\mathbf{E}_e^{(q)}$ is used, and is given as:

$$\mathbf{E}_e^{(q)} \cong \mathbf{E}^{(q)} + \Delta \overline{\mathbf{E}}^{(q)} \tag{4.12}$$

where $\mathbf{E}^{(q)}$ is exactly the field obtained before the order was increased, and $\Delta \overline{\mathbf{E}}^{(q)}$ is a linear combination of the new basis functions added, i.e., an approximation to $\Delta \mathbf{E}^{(q)}$

from (4.8). To find $\Delta\overline{\mathbf{E}}^{(q)}$, (4.12) is substituted into (4.11), and required to hold for all weight functions $\Delta\mathbf{w}$ that are linear combinations of the new basis functions. This results in the expression:

$$a(\Delta\overline{\mathbf{E}}^{(q)}, \Delta\mathbf{w}) = c^{(q)}(\Delta\mathbf{w}) \tag{4.13}$$

which is essentially (4.11) with the old degrees of freedom frozen (i.e., unchanged). Equation (4.13) is a small problem involving only element e and its neighbors, and is very cheap to solve.

This leads to an expression for the Scheme 2 error indicator as follows:

$$\Delta S_{pq} = \left| c^{(p)}(\Delta\overline{\mathbf{E}}^{(q)}) \right| \tag{4.14}$$

The expression in (4.14) provides a close approximation to the exact error indicator given by (4.8). The advantage of this formulation is that it is independent of changes in the original degrees of freedom; any inaccuracy in the approximation is dependent only on the inaccuracy in the highest order coefficients of the element. This results in a relatively accurate yet inexpensive method to calculate $\Delta\overline{\mathbf{E}}^{(q)}$, considering just one element at a time and solving a local problem.

4.3.1.2 Implementation

The implementation of the error indicator given by (4.14) is fairly simple. Once the global matrix $[A]$ and right hand side vector $\{b\}$ is built with elements one order higher, they are partitioned in to parts that are related to the original basis function and the parts related to the new basis functions as follows:

$$[A] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \; ; \; \{b^{(q)}\} = \begin{Bmatrix} b_1^{(q)} \\ b_2^{(q)} \end{Bmatrix} \tag{4.15}$$

From this and the solution vectors $\{E^{(q)}\}$, the column vector $\{c_2^{(p)}\}$associated with the linear form $c^{(p)}$ in (4.9) can be constructed as follows:

$$\left\{c_2^{(p)}\right\} = \left\{b_2^{(p)}\right\} - \left[A_{21}\right]\left\{E^{(q)}\right\} \tag{4.16}$$

To obtain the approximation to the error indicator as given by (4.14) for each element e, the following procedure is used: a submatrix $[A_{22e}]$ and subvectors $\left\{c_{2e}^{(p)}\right\}$ are found, corresponding to the new basis functions for just this element. To find the field due to the new basis functions, $\Delta\overline{\mathbf{E}}^{(q)}$, the following expression is evaluated:

$$\left[A_{22e}\right]\left\{\Delta\overline{E}^{(q)}\right\} = \left\{c_{2e}^{(p)}\right\} \tag{4.17}$$

This expression is essentially (4.13) written in matrix form. Once $\Delta\overline{\mathbf{E}}^{(q)}$ is obtained, the approximation to the error indicator given by (4.14) can be found as follows:

$$\Delta S_{pq} = \left|\left\{c_{2e}^{(p)}\right\}^T \left\{\Delta\overline{E}^{(q)}\right\}\right| \tag{4.18}$$

where the superscript T denotes the transpose.

4.3.2 Error indicator results

The new Scheme 2 error indicator was tested on a number of 3D microwave problems, and compared to the Scheme 1 indicator mentioned previously. For the adaptive FE analysis, the hierarchal, vector tetrahedral elements outlined in Section 4.2 were utilized. The return loss ($20\log_{10}|S_{11}|$) was targeted for all the test cases. At each adaptive step, the error indicator was used to determine the error level in each element with respect to the targeted parameter, and the top 25% of the elements were chosen for refinement. The difference between the resulting return loss value (the 'adaptive result') and the return loss value obtained with all elements set to Order 3 (the 'reference result') was designated as the overall return loss error for each scheme, and plotted against the cumulative computational cost. The return loss value obtained with uniform Order 3 elements was presumed to be the most accurate for the given mesh, and hence used as the reference.

The cumulative cost was obtained by adding the costs of the current and all preceding adaptive steps, with the cost per step being the number of floating point operations (FLOPs) needed to solve the global matrix using a pre-conditioned conjugate gradient solver. The cost at each step is computed using the number of non-zero elements in the sparse global matrix, and the number of steps the pre-conditioned solver takes to solve the matrix system. Although this does not account for the total computational effort required at each analysis point, it provides a good representation of the total cost.

A U-bend rectangular waveguide, a waveguide transformer and a two-cavity iris filter were selected as test cases for the Scheme 2 error indicator. The change in the return loss error values were tracked for both Scheme 1 and Scheme 2 indicators over several adaptive steps, and plotted. The results were reported in [85].

4.3.2.1 U-bend

The first device is an E-plane U-bend shown in Figure 3.7. The design parameter values are given by $\{g\}^T = \{10.659, 14.826, 3.961, 3.048\}$mm, and the number of tetrahedral elements in the mesh is 1471. The separation between the arms (controlled by g_4) and the miter bend angle (controlled by g_1, g_2 and g_3) influence the performance of the device. The operating frequency range is 10-15 GHz. The same mesh was used for both Scheme 1 and Scheme 2 analysis.

The frequency response of the device using the three uniform element orders (Orders 1, 2 and 3) are given in Figure 4.1. Although the adaptive analysis is done only at a single frequency point, the response curves illustrate the changes in the overall frequency range performance of the device as the mesh becomes more refined. From the figure it is obvious that for Order 1 the device is under-discretized, while Order 2 and 3 provide more accurate responses, with Order 3 presumed to be the most accurate.

The adaptive frequency is selected at 12.5 GHz, and the results of the adaption using both error indicator schemes are shown in Figure 4.2. Although neither one gives a monotonic decrease in the error, Scheme 2 manages to give significantly lower error than

Scheme 1 for the same computational cost. The under-discretization due to low order elements seems to affect the performance of the Scheme 1 indicator, with the error values rising considerably in the early stages before being reduced to 0, even though both Scheme 1 and 2 start off with similar error values.

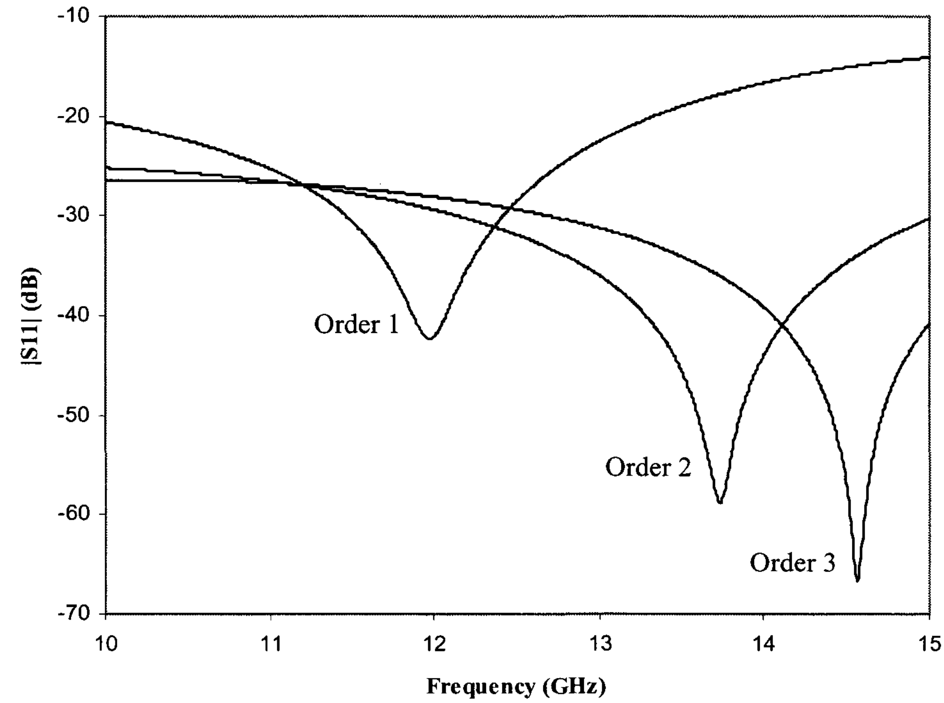

Figure 4.1 Return loss at either port of the U-bend for three uniform orders, as a function of frequency

4.3.2.2 Waveguide impedance transformer

The second example is the waveguide impedance transformer shown in Figure 3.10, with an operating frequency range of 9.75 GHz-10.25 GHz. The performance of the device is controlled by the dimensions of the middle section (a, b, l). The return loss at the larger port (Port 1: $a_1 \mathrm{x} b_1$) is targeted for the analysis. A fairly refined mesh with 1274 tetrahedral elements is used for the analysis. Figure 4.3 shows the frequency response obtained for the given design configuration using the three uniform orders.

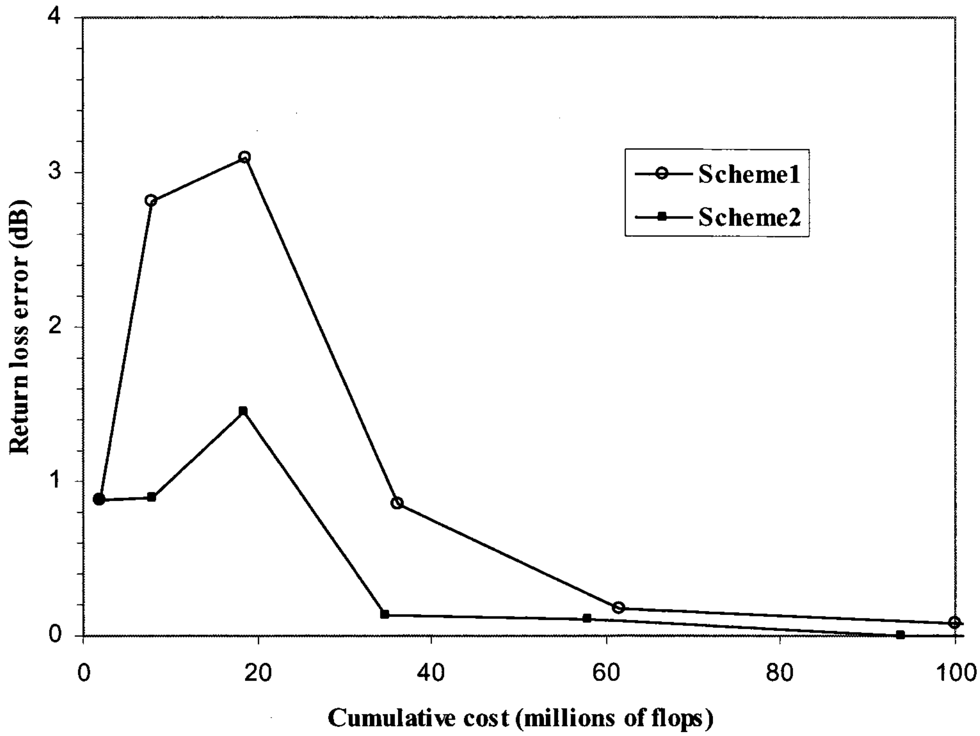

Figure 4.2 Return loss error at either port of the U-bend, for two adaptive schemes. Adaptive frequency=12.5 GHz, No. of tetrahedra=1471

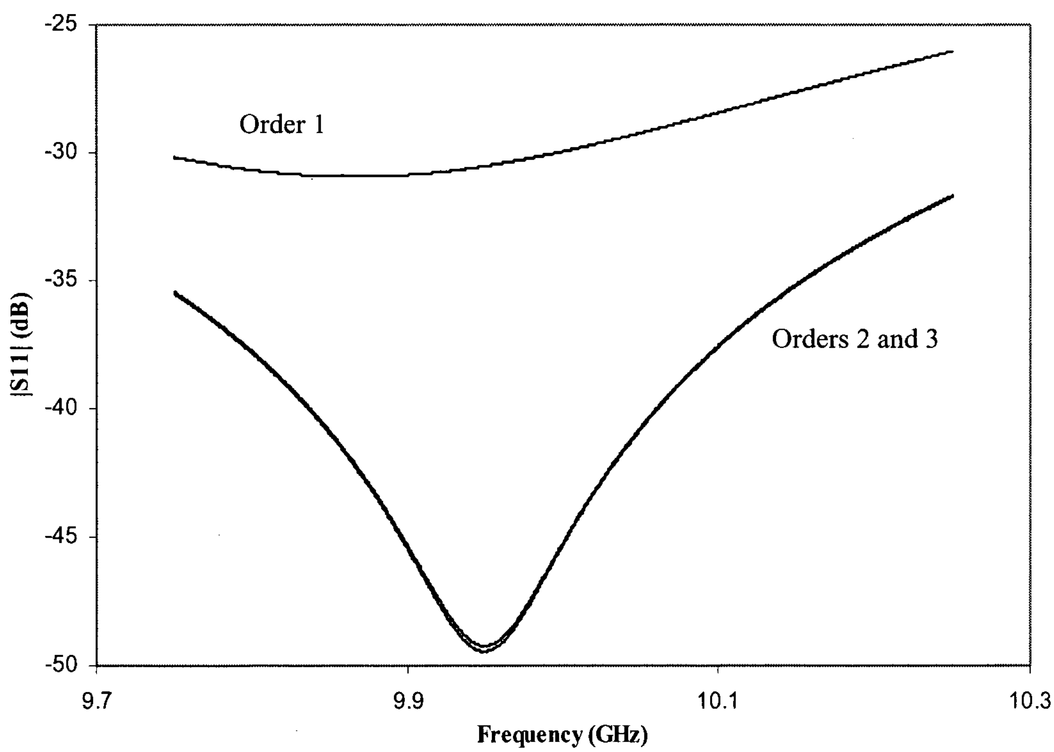

Figure 4.3 Return loss at the larger port (a_1 x b_1) of the waveguide transformer

Although the plots for Orders 2 and 3 are similar, the Order 1 curve shows significant error. The adaptive frequency is selected as 10GHz, and the error reduction by the adaptive analysis is shown in Figure 4.4. The considerable error in the FE solution due to the under-refinement by using uniform Order 1 elements is evident in Figure 4.4 as well, where the initial error for both schemes starts from a fairly high value of approximately 16 dB, but is quickly reduced in subsequent adaptive steps. The analysis using Scheme 2 is able to bring the error values down close to zero within four-five adaptive steps: approximately half of what Scheme 1 required.

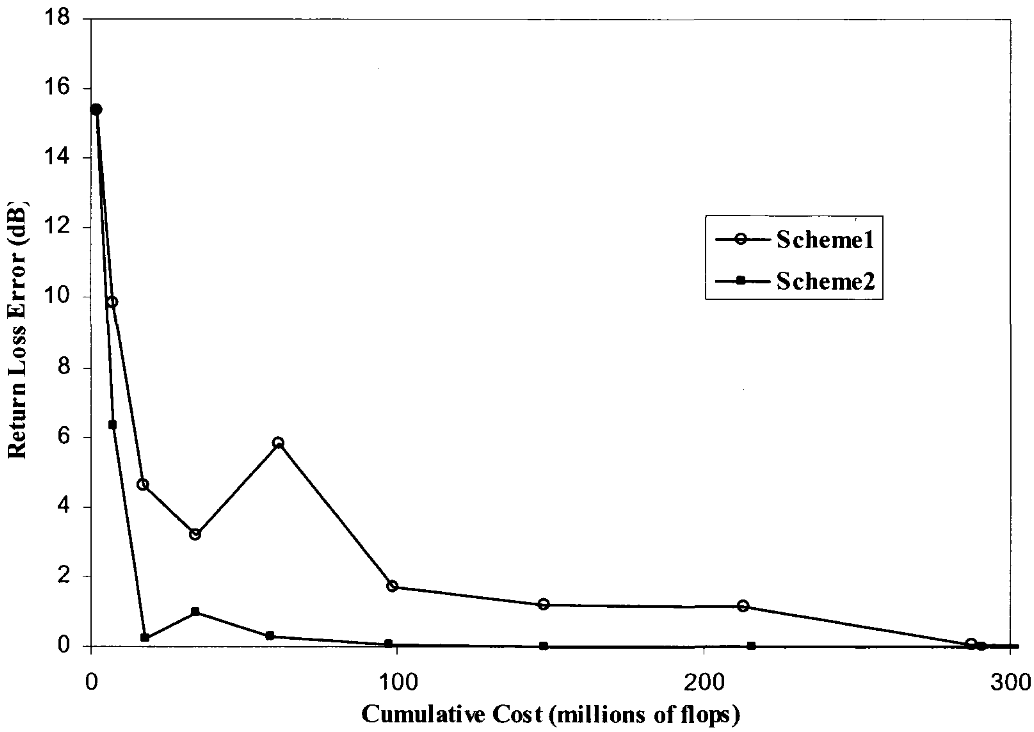

Figure 4.4 Return Loss Error for the larger port (a_1 x b_1) of a waveguide transformer for two adaptive schemes. Adaptive frequency=10 GHz; No. of tetrahedra = 1274.

4.3.2.3 Two-cavity iris filter

The final test case for the error indicator is the two-cavity iris filter shown in Figure 4.5. The performance of the device is controlled by the cavity lengths between irises as well as the iris window dimensions and positioning. The dimensional details are given in [86]. A well-refined mesh of 3767 tetrahedral elements is used for the analysis in order to properly model this complex device. The computed frequency response of the device for

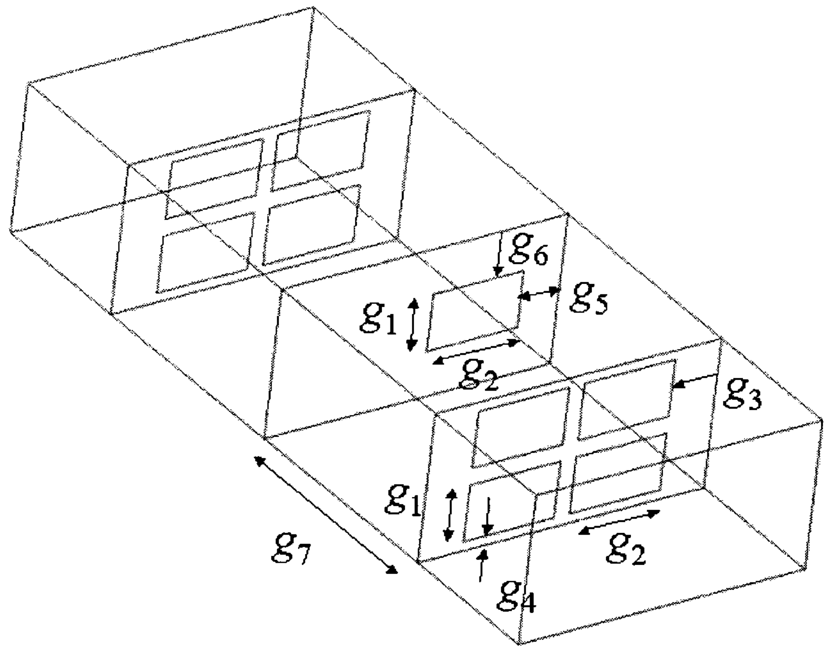

Figure 4.5 Two-cavity iris filter in rectangular waveguide. All metal surfaces are approximated with a PEC boundary condition; $\{g\}^T$ = {3.0, 5.0, 2.47, 0.5, 2.133, 2.45, 12.155} mm; Adaptive analysis frequency = 14.6 GHz; No. of tetrahedra = 3767. [86]

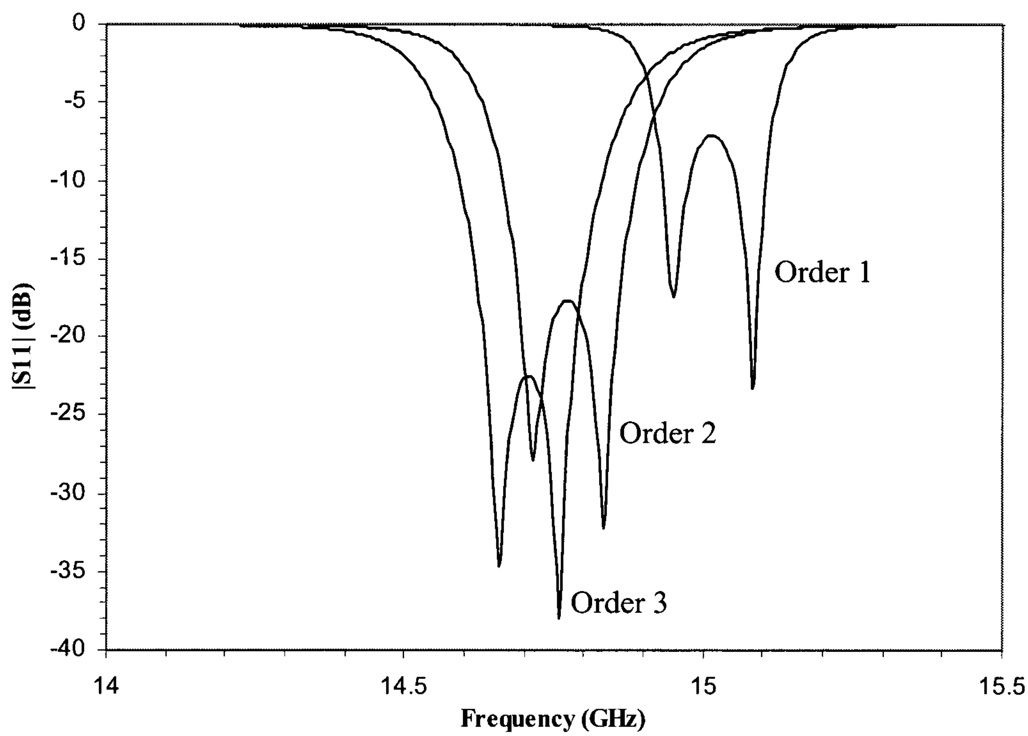

Figure 4.6 Return loss for the two-cavity iris filter at either port

the given design configuration using the three uniform orders is shown in Figure 4.6. Although the Order 1 curve is slightly removed from the other two response curves, the frequency shift between the Order 2 and 3 responses is less than 0.1 GHz, which is indicative of high degree of convergence.

Figure 4.7 shows the return loss error reduction by the adaptive analysis for both error indicators. Both methods manage to reduce the error considerably, although Scheme 2 does so much more quickly: the error value is reduced to 1 dB by Scheme 2 at approximately half the cost required by Scheme 1, which is also evident from the plot.

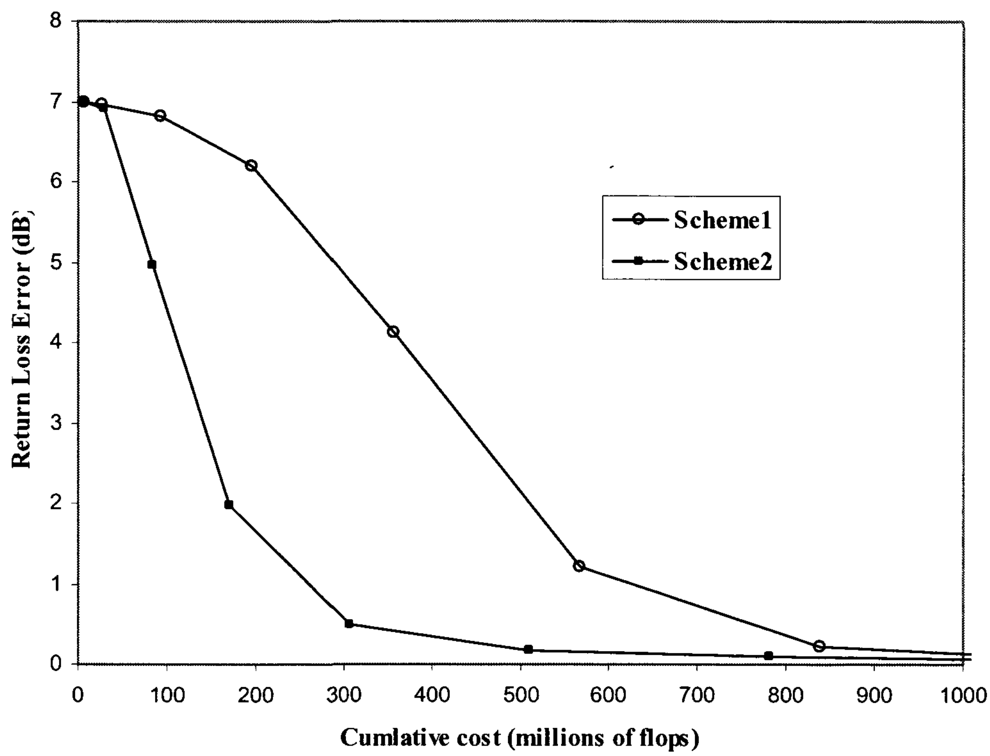

Figure 4.7 Return Loss error for the filter, for two adaptive schemes. Adaptive frequency=14.6 GHz, No. of tetrahedra=3767

Based on the three test cases used here to illustrate the new targeted error indicator, it is clear that Scheme 2 is able to outperform the Scheme 1 approach on a consistent basis. For all three test cases, of varying complexity, from the simple U-bend to the complicated two-cavity filter, the Scheme 2 adaptive analysis is able to reduce the error levels in the return loss significantly at low computational costs. This establishes the effectiveness of the new forward-looking error indicator for the changing scattering parameter values.

4.4 Error Estimation

An accurate error estimate for the quantity of interest is critical in determining the appropriate termination point of the adaptive process. While the error indicator discussed in the previous section evaluates the error on a relative basis, the error estimator gives an assessment of the overall accuracy of the solution obtained from the refined mesh, determining whether the required accuracy has been attained or not [58]. The estimator needs to compute an error that is not too different from the actual error in order to prevent the adaptive process from terminating prematurely (which would give inaccurate solutions), or allowing it to go further than required (resulting in higher computational costs). A brief survey of *a-posteriori* error estimators used in the finite element analysis has been provided in [87], [88].

For gradient based optimization techniques, the availability of accurate sensitivity values is critical to the optimization. These sensitivity values, which are computed from the field solution of the FE analysis, are subject to discretization error. Hence a method to compute accurate *a-posteriori* error estimates targeting design sensitivities is crucial for the design optimization method outlined in this thesis.

There has been relatively little published on error estimation for global quantities extracted from FE solutions, and almost nothing on sensitivity errors. Previous work in this area have produced estimation schemes that provide a relative measure of the error for each finite element, similar to the error indicators defined in this thesis, for example [80], [89],[90]. However they all fail to give an absolute measure of the error in the quantity of interest. An alternative is to use dual or complementary analyses, where the error is the difference between two solutions of the same global quantity obtained from problems of roughly the same size [91], [92].

In [93], a more targeted formulation for estimating errors in quantities of interest using measures other than the energy norm was proposed. Another such approach for estimating scattering parameter sensitivity errors for microwave devices was formulated in [94]. However, this approach lacked a strong theoretical foundation, and it was dependent on the number and placement of internal nodes in the FE mesh.

In more recent work, a general approach to calculating error estimates for functionals of the field using adjoint solutions, known as the Dual Weighted Residual (DWR) method, was proposed [95]. The same work also provided a survey of error estimators for specific quantities of interest using the duality principle. The DWR method was applied to estimate errors in scattering parameters in [96], and to some non-linear electrostatic quantities (force, capacitance) in [97]. For this thesis, the theory was extended to calculate the error estimates of the scattering parameter sensitivities.

4.4.2 Error Estimator theory

The general theory for estimating the error in the design sensitivities is presented first, followed by the modifications made for calculating the error in scattering parameter sensitivities.

Suppose that an unknown field **E** can be found using the FE analysis with the following equation:

$$a(\mathbf{E},\mathbf{w}) = b(\mathbf{w}) \quad \forall \mathbf{w} \in H \tag{4.19}$$

where a is the bilinear form, b is a linear functional and H is the space spanned by the FE basis functions. Suppose $G(\mathbf{E})$ is a quantity of interest, a functional of the field, e.g. a scattering parameter. The sensitivity of G can be obtained inexpensively, once the adjoint solution $\mathbf{E}_a$ is found by solving:

$$a(\mathbf{E}_a,\mathbf{w}) = \delta G(\mathbf{E};\mathbf{w}) \quad \forall \mathbf{w} \in H \tag{4.20}$$

where $\delta G(\mathbf{E};\mathbf{w})$ is the first variation of G. Using the adjoint solution $\mathbf{E}_a$, and the FE solution $\mathbf{E}$, the expression for the sensitivity of G with respect to any design parameter g is given as:

$$\left.\frac{dG}{dg}\right|_{FE} = \frac{dG}{dg}(\mathbf{E},\mathbf{E}_a) = \frac{\partial G}{\partial g}(\mathbf{E}) + \frac{\partial b}{\partial g}(\mathbf{E}_a) - \frac{\partial a}{\partial g}(\mathbf{E},\mathbf{E}_a) \tag{4.21}$$

The first step to obtain an error estimate is to see how much $\mathbf{E}$ would change when the polynomial order of *all* hierarchal FEs are increased by one. Part of the change, $\Delta\overline{\mathbf{E}}$, is due to the new basis functions added and can be obtained approximately by solving the following linear equation:

$$a(\Delta\overline{\mathbf{E}},\Delta\mathbf{w}) = b(\Delta\mathbf{w}) - a(\mathbf{E},\Delta\mathbf{w}) \tag{4.22}$$

for all $\Delta\mathbf{w}$ that are linear combinations of the new basis functions added. The underlying concept here is similar to what was presented in the error indicator section, with (4.22) being similar to (4.13), where the formulation targeted the change in the field for a single element. The linear system in (4.22) can be solved inexpensively by ignoring off-diagonal entries in the matrix corresponding to the left hand side. The estimate $\Delta\overline{\mathbf{E}}$ can be further improved by adding to it a field $\mathbf{e}$ in H, found by solving:

$$a(\mathbf{e},\mathbf{w}) = -a(\Delta\overline{\mathbf{E}},\mathbf{w}) \qquad \forall \mathbf{w} \in H \tag{4.23}$$

Adding the field $\mathbf{e}$ removes any component from $\Delta\overline{\mathbf{E}}$ that lies in the space H that is spanned by the *original* basis functions, thus improving the accuracy of the estimate. This calculation needs to be done only once, as was shown in [97], regardless of the number of design parameters.

A similar approach is used to find the corresponding quantities $\Delta\overline{\mathbf{E}}_a$ and $\mathbf{e}_a$ for the adjoint solution, resulting in expressions similar to (4.22) and (4.23). The expression for the approximate change in the adjoint solution can be written as:

$$a(\Delta\overline{\mathbf{E}}_a,\Delta\mathbf{w}) = \delta G(\mathbf{E}+\mathbf{e}+\Delta\overline{\mathbf{E}};\Delta\mathbf{w}) - a(\mathbf{E}_a,\Delta\mathbf{w}) \tag{4.24}$$

and the field $\mathbf{e}_a$ in H is obtained as the solution to:

$$a(\mathbf{e}_a,\mathbf{w}) = \delta G(\mathbf{E}+\mathbf{e}+\Delta\overline{\mathbf{E}};\mathbf{w}) - \delta G(\mathbf{E};\mathbf{w}) - a(\Delta\overline{\mathbf{E}}_a,\mathbf{w}) \tag{4.25}$$

An improved value of the design sensitivity value can be obtained as:

$$\left.\frac{dG}{dg}\right|_{improved} = \frac{dG}{dg}(\mathbf{E} + \mathbf{e} + \Delta\overline{\mathbf{E}}, \mathbf{E}_a + \mathbf{e}_a + \Delta\overline{\mathbf{E}}_a) \tag{4.26}$$

The error estimate for the design sensitivity value is then the difference between the improved value and the original sensitivity value:

$$\Delta\overline{G} = \left.\frac{dG}{dg}\right|_{improved} - \left.\frac{dG}{dg}\right|_{FE} \tag{4.27}$$

The computational effort required to find the error estimate values is associated with computing the field values $\mathbf{e}$ and $\mathbf{e}_a$, and is about the same as that required to find the sensitivity itself using the adjoint. However, the advantage is that these steps are independent of the number of geometric parameters. A direct application of the DWR method outlined in [95] would have required the computation of a different adjoint solution for each design parameter. The alternative scheme presented in [97] avoids the use of the adjoint solution, and is used here to calculate the values for $\mathbf{e}$ and $\mathbf{e}_a$.

The general theory presented above can now be applied to the specific case of the scattering parameters S_{pq} of an N-port microwave device obtained from $\mathbf{E}^{(q)}$, the electric field when port q is excited with all the other ports matched and given by (4.3). The quantity of interest for the error estimation formulation used here becomes:

$$G := S_{pq} = b^{(p)}(\mathbf{E}^{(q)}) - \delta_{pq} \tag{4.28}$$

Let the field components $\Delta\overline{\mathbf{E}}$ and $\mathbf{e}$ coming from $\mathbf{E}^{(q)}$ be denoted $\Delta\overline{\mathbf{E}}^{(q)}$ and $\mathbf{e}^{(q)}$, respectively. Then using (4.28), it can be shown that the corresponding quantities for the adjoint solution are just those that come from $\mathbf{E}^{(p)}$, i.e.,

$$\mathbf{E}_a = \mathbf{E}^{(p)};\ \ \Delta\overline{\mathbf{E}}_a = \Delta\overline{\mathbf{E}}^{(p)};\ \ \mathbf{e}_a = \mathbf{e}^{(p)} \tag{4.29}$$

This results in a situation where once the values of $\mathbf{E}$, $\Delta\overline{\mathbf{E}}$ and the correction value $\mathbf{e}$ are computed, the calculations for the corresponding adjoint components can be avoided, resulting in further computational savings.

4.4.3 Error estimation results

The effectiveness of the new error estimation theory was tested on three 3D microwave devices. The new estimator was combined with the Scheme 2 error indicator, outlined earlier in this chapter, to form a true p-adaptive scheme. The analysis was run for 10 adaptive steps for each of the examples. At each adaptive step, 25% of the elements with the highest error (as per the Scheme 2 indicator) were refined. The resulting mesh was used to find the scattering parameter values and their sensitivities, as well as the error estimates. These estimates were compared to the 'true error' values, which were obtained by taking the difference between the FE analysis results for a given adaptive step, and a converged, reference value. The reference value was the sensitivity value obtained by using the highest element order possible of the hierarchal basis functions (Order 3 for the current case) and a very fine mesh.

A two-cavity band-pass filter, a waveguide transformer and a magic-T were used to test the new error estimation formulation. The results were reported in [98].

4.4.3.1 Two-cavity iris filter

The first test case is the two-cavity filter shown in Figure 4.5, with all the design parameters indicated in the figure. The reflection coefficient S_{11}, measured at either port (due to symmetry) was targeted. The performance is governed by the seven geometric parameters that control the length of the two cavities, as well as the length and width of the iris windows. The iris thickness is fixed, as well as the waveguide width and height. The error estimates for the sensitivities of $G=S_{11}$ were found with respect to all the design parameters and compared to the true value. The estimated error and the true error are given by:

$$\text{Estimated} = \frac{\left|\Delta \overline{G}\right|}{\left|(dG/dg)_{ref}\right|} \times 100\% \tag{4.29}$$

$$\text{True} = \frac{\left|(dG/dg)_{FE} - (dG/dg)_{ref}\right|}{\left|(dG/dg)_{ref}\right|} \tag{4.30}$$

Here $(dG/dg)_{ref}$ is the reference value, $(dG/dg)_{FE}$ is the value obtained at a given adaptive step, and $\Delta\overline{G}$ is the error estimate given by (4.27). A fairly refined mesh with 3821 tetrahedral elements was used for the adaptive analysis. The converged values for the reference were obtained using a mesh with approximately 16,000 uniform Order 3 elements.

The graphs of the estimated error plotted against the true error for the sensitivity with respect to design parameter g_5 is shown in Figure 4.8. Each square marker corresponds to one adaptive solution in the FE analysis. The square at the top right represents the solution obtained with all the elements set to Order 1, and as the adaption progresses, some elements are increased in order, which results in a decrease in the error.

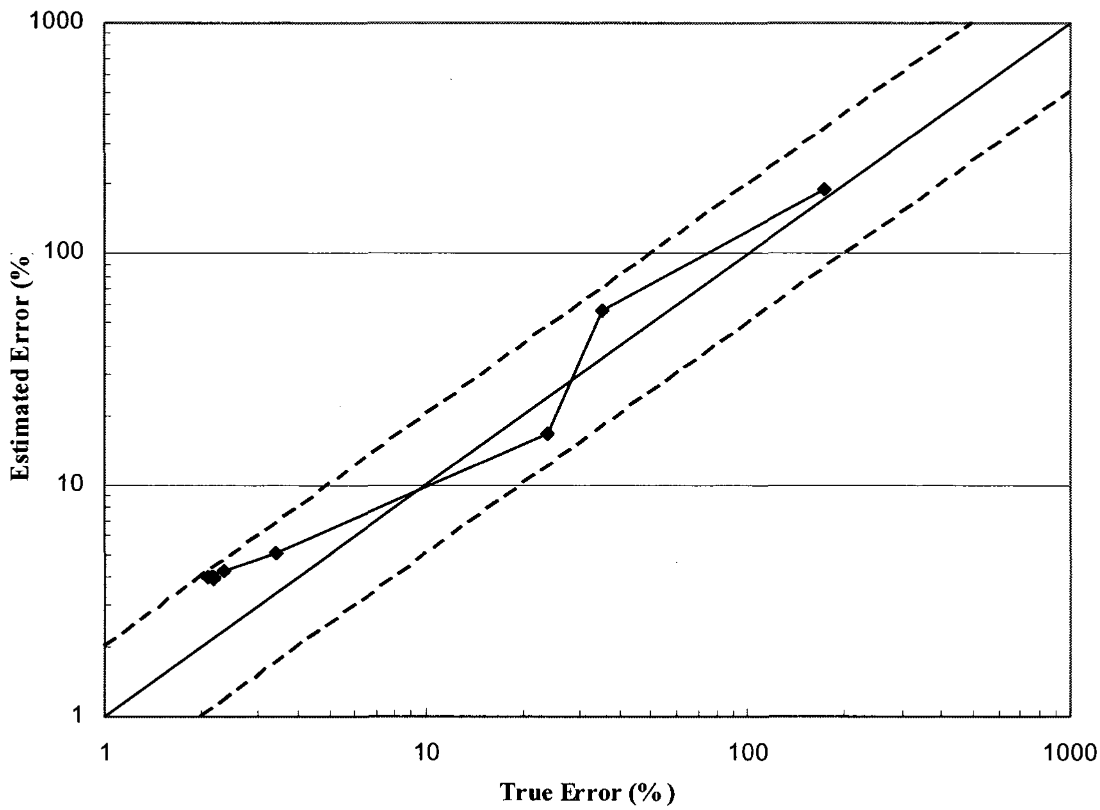

Figure 4.8 Error in the sensitivity of S_{11} with respect to parameter g_5 shown in Figure 4.5. Adaptive frequency = 15GHz; No. of tetrahedra=3821

The above plot shows that the estimator is able to give good estimates of the errors over a wide range spanning several adaptive steps, and within a 2-to-1 ratio factor (represented by the dashed diagonal lines) with respect to the true error. In the late stages of the analysis, the estimator tends to over-estimate the error values, with ratio values greater than one. From an adaptive analysis perspective, slight over-estimation of the error value might be preferred, as under-estimating the error could result in the analysis being terminated prematurely, and consequently, less accurate solutions.

The next plot in Figure 4.9 shows the error estimate for the sensitivity values with respect to the critical design parameter g_7. This parameter controls the cavity length of both cavities of the filter, and to which the device performance is highly sensitive. The length of both cavities is controlled by the same design parameter, thus maintaining the overall symmetry of the device.

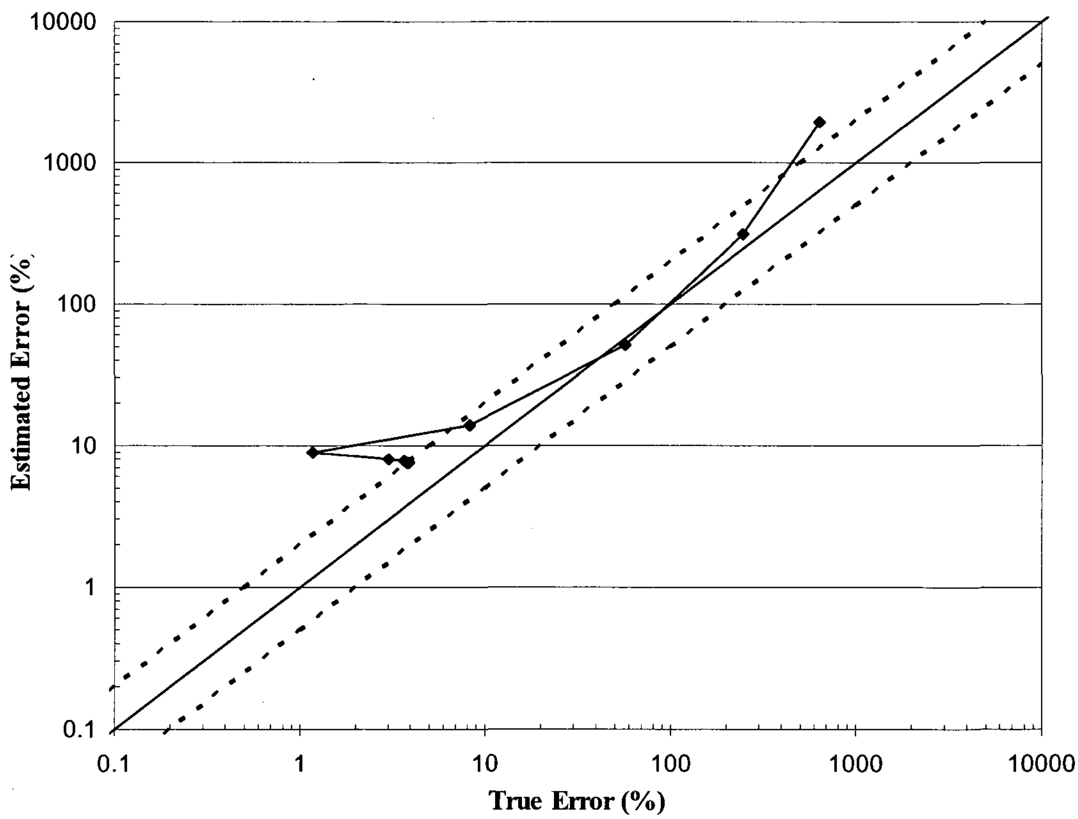

Figure 4.9 Error in sensitivity with respect to g_7 shown in Figure 4.5. Adaptive frequency=15GHz, No. of tetrahedra=3821

The plot demonstrates the error estimator's ability to produce sensitivity error estimates reasonably close to the true error at various stages of the adaptive analysis: most values are more or less within the 2-to-1 ratio band shown in the figure, with the exception of the value at the 5th adaptive step. As the adaption progresses beyond this point, the estimated-to-true error value ratios converge to between 1-10%, which is acceptable. The point that corresponds to the 5th adaptive step that shows drastic over-estimation could be a spurious point during the analysis. Beyond this point, the analysis is able to bring the error ratios close to the band, with over-estimation by a factor of approximately two.

At a given adaptive step, the analysis provides the error estimates for all 7 sensitivities, and Figure 4.10 shows the estimate values compared with the true error values. Each diamond-shaped marker represents the estimate-to-true value ratio for each geometric parameter obtained from one analysis point (the fifth adaptive step). The plot demonstrates the wide range of errors: while most of the sensitivity errors are less than 10%, there are two sensitivities with substantial error. This illustrates the critical role of the error estimator in identifying the sensitivities which have yet to attain acceptable accuracy levels, and hence unsuited to be passed on to the optimizer.

Another convenient way of interpreting the estimator results for each adaptive step is to look at the sensitivity vector $\{s\}$, defined by:

$$\{s\} := \{\partial G/\partial g_1, \partial G/\partial g_2, \ldots \partial G/\partial g_n\}^T \tag{4.31}$$

where n is the total number of design parameters.

The corresponding vector for the estimated errors is given as:

$$\{\Delta\bar{s}\} := \{\Delta\bar{G}_1, \Delta\bar{G}_2, \ldots \Delta\bar{G}_n\}^T \tag{4.32}$$

At each adaptive step, the estimated and true errors can then be defined using the infinity norms of the vectors as follows:

$$\text{Estimated} = \frac{\left\|\{\Delta\bar{s}\}\right\|_\infty}{\left\|\{s\}_{ref}\right\|_\infty} \times 100\% \tag{4.33}$$

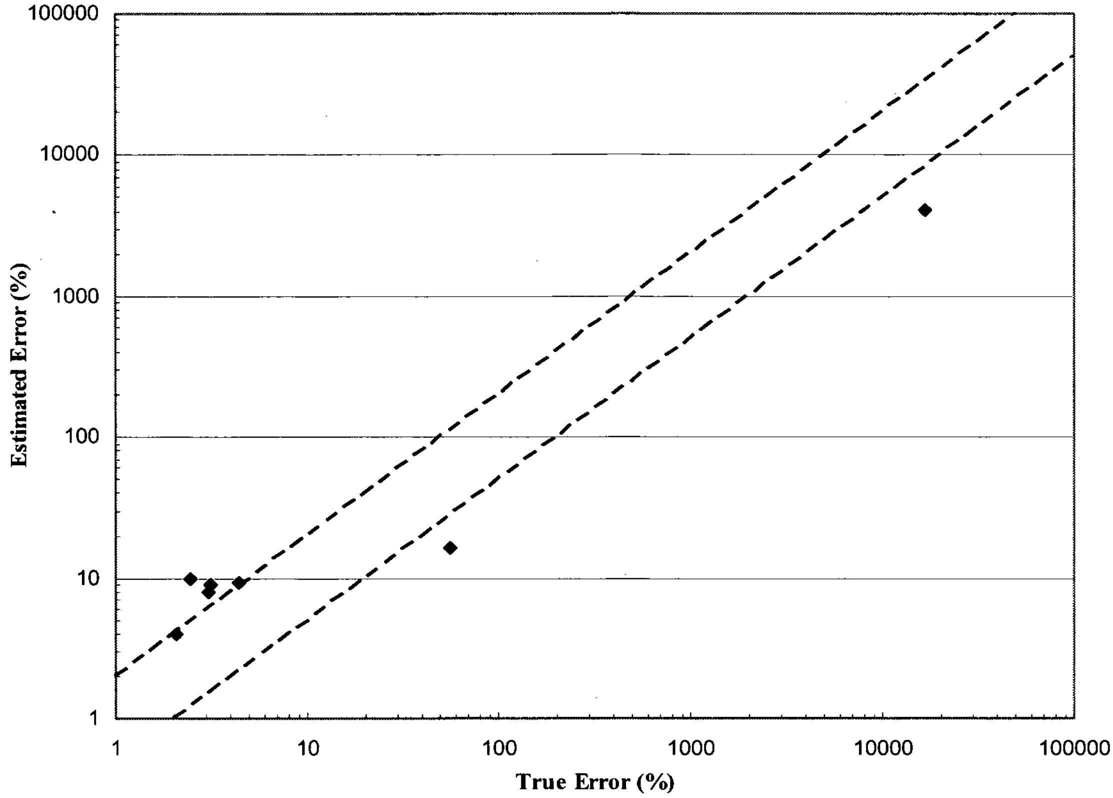

Figure 4.10 Error in the sensitivity of S_{11} to each of the 7 geometric parameters shown in Figure 4.5. Each point corresponds to one parameter, and all points are for one analysis, at 15 GHz. No. of tetrahedra=3821

$$\text{True} = \frac{\left\| \{s\}_{FE} - \{s\}_{ref} \right\|_\infty}{\left\| \{s\}_{ref} \right\|_\infty} \times 100\% \tag{4.34}$$

Thus for each step, this error estimator formulation tries to characterize the sensitivity error using the worst value for a given error vector, and the adaptive scheme tries to attain the desired solution accuracy by reducing these values.

The results obtained from the adaptive scheme applied to the next two examples use (4.33) and (4.34) as the estimated and true error values.

4.4.3.2 Waveguide transformer

The next test case is the waveguide transformer shown in Figure 3.10. The return loss at the larger port was targeted, and the 25% adaptive scheme was used for mesh refinement. A relatively coarse mesh with 229 elements was used for the adaptive analysis, while the reference values were obtained using a mesh with approximately 7000 tetrahedra and uniform Order 3 elements. The resulting plot is shown in Figure 4.11.

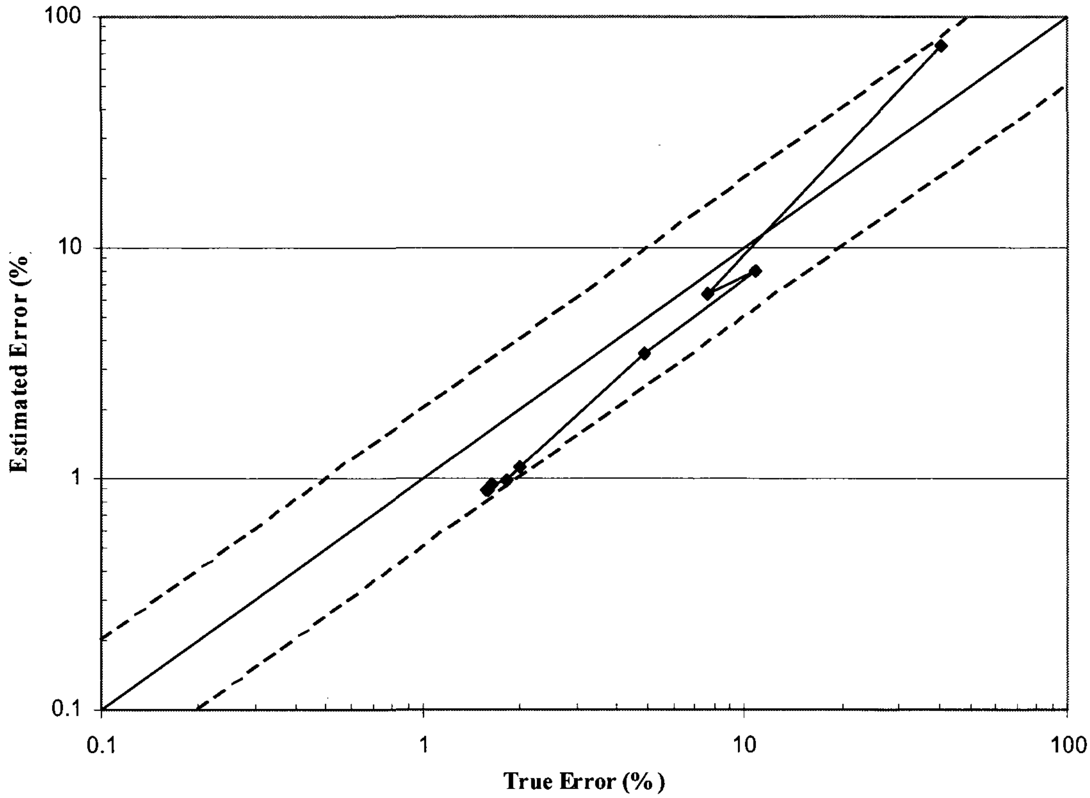

Figure 4.11 Error in the sensitivity vector for the waveguide transformer during adaptive analysis. Adaptive frequency=10GHz, No. of tetrahedra=229

Each marker indicates the ratio values between the estimated and true error values obtained using the formulation given by (4.33)-(4.34) at each adaptive step. The adaptive scheme is able to bring down the true error from around 100% to as low as approximately 1%. The estimator is able to produce error estimates acceptably close to the true error values, with all the points falling within the 2-to-1 ratio band. Although the results show under-estimation towards the end, these happen for estimated error values less than 5%.

4.4.3.3 Magic-T

The final example is a compensated magic-T shown in Figure 4.12; dimensional details are given in [4]. A function G which measures how well the device is matched at two ports was used for the sensitivity error calculations. The function is:

$$G = |S_{11}|^2 + |S_{44}|^2 \tag{4.35}$$

To find the sensitivities, (4.35) was directly differentiated with respect to the design parameters, and then the regular FE and improved values for S_{11}, $\partial S_{11}/\partial g$, S_{44} and $\partial S_{44}/\partial g$ were used to compute the error estimate given by (4.26).

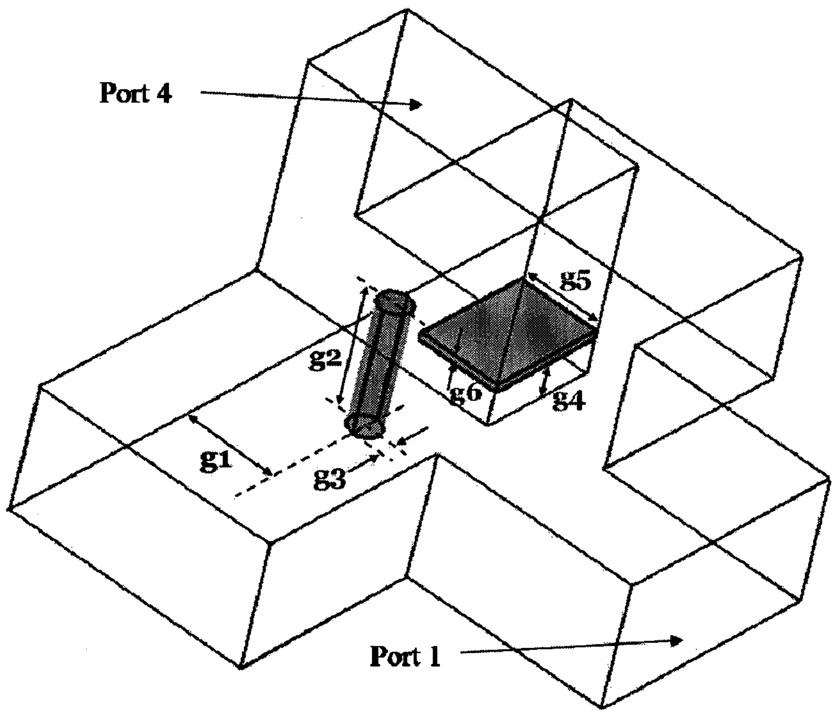

Figure 4.12 Compensated Magic-T. All four waveguide ports are 25.4x12.7mm. $\{g\}^T$ ={9.8125, 12.7, 1.5875, 3.556, 8.382, 0.7939} mm; [4].

The performance of the device is dependent on the positioning and dimensions of the post and the iris, which in turn, are controlled by the 7 geometric parameters indicated in the figure. The sensitivity values with respect to all the design parameters and their error estimates were calculated for the adaptive analysis. The reference value for the true error calculation was obtained by using a very fine mesh with approximately 20,000 tetrahedral elements, while the regular FE adaptive analysis was done with a mesh of 9250 elements. The resulting plot over eight adaptive steps is given in Figure 4.13.

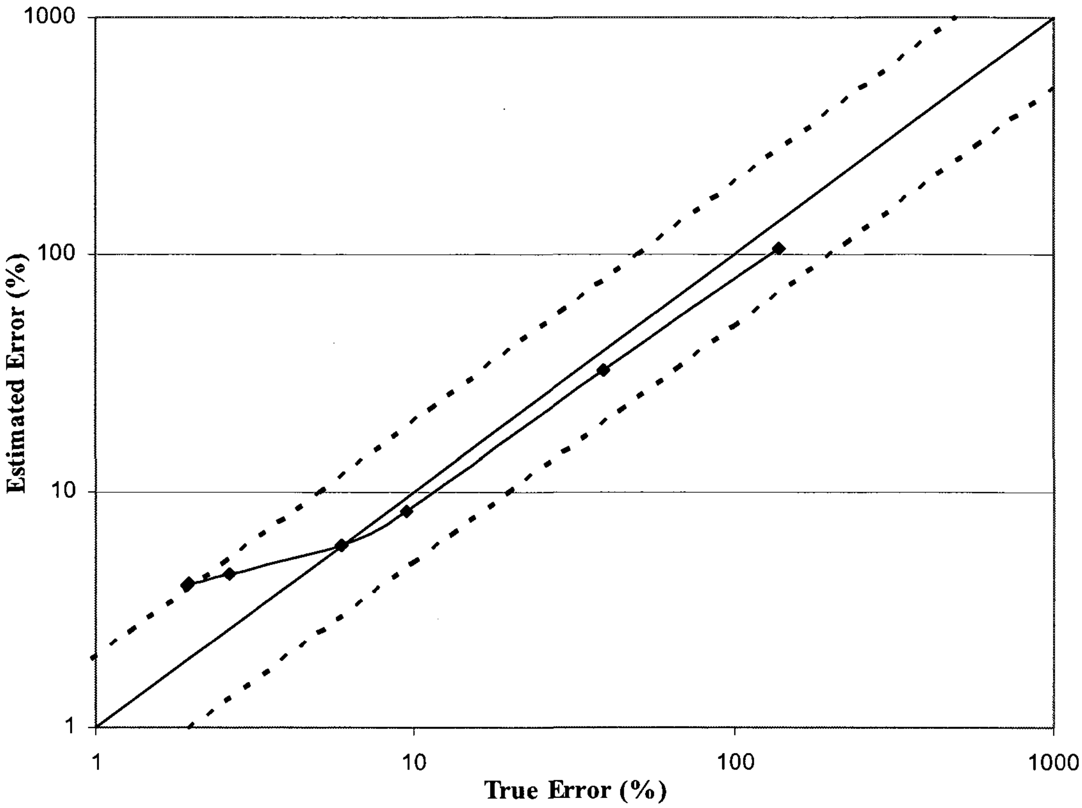

Figure 4.13 Error in the sensitivity vector for the magic-T obtained using adaptive analysis. Analysis frequency=9 GHz. No. of tetrahedra=9249

The graph again demonstrates the estimator's ability to produce error estimates close enough to the true value throughout the analysis, with all the values lying within the 2-to-1 ratio band. The estimator is able to produce accurate estimates even in the early stages of the analysis where low order elements are used. As the analysis progresses, the estimator is able to produce more accurate estimates (compared to the true error values), coming closer to the unity-ratio line. The final stages also show the analysis over-estimating the error values. As was explained for the filter example, this is preferred in order to ensure the accuracy of the final gradient values that will be used by the optimizer.

4.5 Discussion

The results shown in this chapter prove the effectiveness of the new error indicator and estimator schemes devised for the *p*-adaptive analysis of 3D microwave problems. The targeted, forward-looking indicator scheme is able to reduce the scattering parameter errors cheaply, while the new estimator is able to produce error estimates for the gradients within acceptable levels consistently, at different stages of the *p*-adaptive analysis that uses the Scheme 2 indicator to guide the refinement.

The combination of the error indicator with the estimator forms the *p*-adaptive analysis that is used for the cost function evaluation module in the optimization process. The efficiency of the overall optimization process is very much dependent on the efficiency with which the CFE module produces results that satisfy the accuracy requirements set by the optimizer. The next chapter investigates how the *p*-adaptive method formulated in this chapter performs when linked with the optimizer in maintaining the efficiency of the optimization analysis.

5. Accuracy controlled optimization of 3D microwave devices

Optimization problems that are solved analytically deal with cost functions that are "infinitely" accurate, and the accuracy of the cost function evaluation process is a non-issue. On the other hand, for practical design optimization problems, the analysis usually needs to rely on numerical approximation techniques. For such cases the accuracy, as well as the computational cost to attain the desired accuracy, is a concern.

Chapter 3 showed how direct optimization techniques can improve microwave device performance by using a cost function and its gradients based on scattering parameters computed using FE. The advantage of direct optimization methods in the context of this research is the relatively small number of analysis steps required to reach a local minimum. The FE analysis is able to produce the gradients in a computationally efficient manner at each optimization step as well. Combining the two features, the optimization scheme can produce significantly improved designs, provided a reasonable initial design configuration is available.

Several researchers have combined numerical techniques (such as FEM, MoM, FDTD etc.) with gradient-based optimization schemes to improve microwave device performance [99]-[102], [42], [103]-[106]. The gradients in these cases were obtained relatively cheaply by avoiding extra solutions using adjoint techniques, thus making the optimization process cost efficient. Despite these gains, solution accuracy and the computational effort involved in the solution process itself at each analysis point warrants special attention for efficiency purposes, especially for 3D microwave problems. From a design optimization perspective, attaching a certain measure of confidence to the accuracy and numerical validity of the solution, while reducing the cost to find it, is clearly desirable.

Conducting the numerical approximation at each optimization step using uniform high order finite elements (as was done in Chapter 3) is computationally expensive, due to the large number of degrees of freedom applied at each CFE. Even though the results obtained are very accurate, the high computational cost at each analysis step leads to cumulative costs that are prohibitive. However, the solution accuracy at various stages of the optimization need not be fixed throughout: the accuracy required at the start of the optimization need not be as stringent as that when closer to the minimum [10]. If the optimizer can relay information regarding the required accuracy level to the FE analysis module for every analysis step, the computation can be performed to varying accuracy levels as the optimization progresses. This can result in considerable computational savings.

Adaptive techniques help achieve these cost savings: the new error indicator and error estimator schemes outlined in Chapter 4 allow for the p-adaptive FE analysis to compute the cost function and its gradients to a specified accuracy level for each analysis step. How this required accuracy level is set by the optimizer with the help of an 'accuracy link' to control the computational costs of the FE analysis is the central idea of this thesis, and is further explored in this chapter.

5.1 The accuracy link between the optimizer and the adaptive FE analysis

The error estimator that is part of the p-adaptive FE analysis module is the key component for the accuracy control mechanism: when the estimated error for the cost function gradient is below a certain level, the adaption (and consequently, the FE analysis) is terminated. The level is set by requiring the relative error in ∇C to be roughly constant throughout the optimization, satisfying the condition:

$$\left\| err\nabla C^{(k)} \right\|_{\infty} < \beta \left\| \nabla C^{(k)} \right\|_{\infty} \tag{5.1}$$

where β is a pre-specified scalar value that determines the accuracy level, and k represents the latest point in design space. Since the cost function gradient tends to zero at the local minimum, this expression will automatically increase the accuracy requirement for the FE analysis as the optimization progresses.

The FE analysis module tries to achieve the condition in (5.1) through p-adaption. The increasing accuracy demands at each consecutive analysis step are met with an increase in the number of DOFs to approximate the solution. The error indicator identifies the high error regions in the mesh for the efficient distribution of the additional

DOFs, and the error estimator gauges the accuracy of the resulting FE approximation to see if the required level was achieved.

The β parameter in (5.1) represents the desired relative error level for the cost function gradient. Setting β too small forces the adaption to use high element orders from the early stages of the analysis, which result in high computational costs. A large value of β will result in low accuracy ∇C values, which will adversely affect the convergence of the optimization. Some experimentation showed that values in the range $0.1< \beta <0.3$ are satisfactory, and a value of β=0.2 (which represents a relative error in $\nabla C^{(k)}$ of 20%) was chosen as a good compromise.

In order to optimize the performance of a given microwave device over a frequency range using the combined adaption-optimization scheme, certain modifications to (5.1) are necessary, as explained in the next section.

5.1.1 Accuracy link formulation for a frequency range

More often, the microwave designer is interested in the device performance over a frequency range rather than for a single frequency. In this case a cost function definition based on the scattering parameter behavior over the range is assumed. Chapter 3 showed how the response of a device at sample points over the frequency range can be obtained using the Pade approximation technique, by calculating the moments of the scattering parameter values and their sensitivities. How the general accuracy link can be modified for a frequency range scenario is demonstrated next.

The cost function values and their gradients at each sample point are computed from the Pade approximation of the scattering parameter value at that point, as explained in Chapter 3. For example, $c_i = |S_{11}|_i$ will be the cost function at the sample point i, obtained using the Pade approximation for S_{11} at that point. Similarly, the cost function gradient vector with respect to the design parameters at the point is calculated as $\nabla c_i = \nabla |S_{11}|_i$.

Consequently, a typical definition for the cost function that represents the device performance over the range can be given as:

$$C = \sum_i c_i \tag{5.2}$$

and the cost function gradients as:

$$\nabla C = \sum_i \nabla c_i \tag{5.3}$$

Based on the definitions of (5.2)-(5.3), the accuracy condition given by (5.1) needs to be adjusted for the frequency range scenario. Following (5.3), the error in the gradient can be given as:

$$\begin{aligned} err\nabla C &= err\sum_i \nabla c_i \\ &= \sum_i err\nabla c_i \end{aligned} \tag{5.4}$$

The gradient error denoted on the right-hand side of (5.4) for a specific i is defined as follows:

$$err\nabla c_i = \nabla c_i\big|_{ref} - \nabla c_i\big|_{FE} \tag{5.5}$$

This exact error definition is applied to all the sampling points in the frequency band to evaluate (5.4).

Applying the norms, the left-hand side of (5.4) can be bound using the triangle inequality [107] as follows:

$$\left\|err\nabla C\right\|_\infty \leq \sum_i \left\|err\nabla c_i\right\|_\infty \tag{5.6}$$

Direct use of (5.6) as a way of estimating the error would require the application of the Pade technique at each adaptive step of the FE analysis. This results in extra solutions to compute the high order moment values of the cost function and its gradients at each step, which is obviously computationally inefficient. It is preferable to conduct the adaptive analysis at one frequency and conduct the frequency sweep only once, after adaption has finished.

The following assumption regarding the error at each sample point is made:

$$\frac{\left\| err\nabla c_i \right\|_\infty}{\left\| \nabla c_i \right\|_\infty} \cong \chi \tag{5.7}$$

where χ is a constant value, holding for all i across the range at a given adaptive step.

Introducing (5.7) into the expression in (5.6), the cost function gradient error value becomes:

$$\begin{aligned} \left\| err\nabla C \right\|_\infty &\leq \sum_i \chi \left\| \nabla c_i \right\|_\infty \\ &= \chi \sum_i \left\| \nabla c_i \right\|_\infty \\ &= \frac{\left\| err\nabla c_a \right\|_\infty}{\left\| \nabla c_a \right\|_\infty} \sum_i \left\| \nabla c_i \right\|_\infty \end{aligned} \tag{5.8}$$

where $\nabla c_a, err\nabla c_a$ are the cost function gradient and the error in the gradient, respectively, computed at the adaption frequency. The error is approximated by the estimate given by (4.27).

From (5.8), it follows that (5.1) will hold provided that:

$$\frac{\left\| err\nabla c_a^{(k)} \right\|_\infty}{\left\| \nabla c_a^{(k)} \right\|_\infty} \sum_i \nabla c_i^{(k)} < \beta \left\| \nabla C^{(k)} \right\|_\infty \tag{5.9}$$

The variable k represents the current design point of the optimization. Equation (5.9) can now be re-arranged to define the accuracy condition for the cost function gradient to be satisfied at each analysis point at the adaptive frequency:

$$\frac{\left\| err\nabla c_a^{(k)} \right\|_\infty}{\left\| \nabla c_a^{(k)} \right\|_\infty} < \beta \frac{\left\| \nabla C^{(k)} \right\|_\infty}{\sum_i \left\| \nabla c_i^{(k)} \right\|_\infty} \tag{5.10}$$

A new term can be introduced to represent the ratio on the right-hand side:

$$\Gamma^{(k)} = \frac{\left\| \nabla C^{(k)} \right\|_\infty}{\sum_i \left\| \nabla c_i^{(k)} \right\|_\infty} \tag{5.11}$$

where $0 \le \Gamma^{(k)} \le 1$, due to the triangle inequality.

The accuracy condition in (5.10) then becomes:

$$\frac{\left\| err \nabla c_a^{(k)} \right\|_\infty}{\left\| \nabla c_a^{(k)} \right\|_\infty} < \beta \, \Gamma^{(k)} \tag{5.12}$$

The strict evaluation of (5.12) would still require the application of Pade approximation at every adaptive step to calculate $\Gamma^{(k)}$. To avoid the extra cost, the ratio term $\Gamma^{(k)}$ is approximated by Γ_{ref}. For example, Γ_{ref} might simply be $\Gamma^{(k\text{-}1)}$ which is available from the previous CFE. That is,

$$\Gamma_{ref} = \frac{\nabla C^{(k-1)}}{\sum_i \left\| \nabla c_i^{(k-1)} \right\|} \tag{5.13}$$

This approximation holds the right-hand side of the accuracy condition in (5.12) constant throughout the adaptive analysis at a given design point.

5.1.2 Implementation of the accuracy link

Figure 3.2 illustrated how the steepest-descent method was implemented using the cubic line search algorithm for the optimization process. As shown, the three main stages of the optimizer use the FE analysis module to calculate the cost function and the gradients. The results in Chapter 3 were obtained non-adaptively using uniform Order 3 elements. Figure 5.1 shows the detailed illustration of the new optimization analysis technique with the new adaptive FE analysis module, whose accuracy requirements are controlled by the optimizer through the accuracy link.

From Figure 5.1, the initial design configuration is used for the pre-optimization analysis to obtain initial values of C and ∇C. The CFE for this step is conducted adaptively to satisfy the condition in (5.12), with $\Gamma_{ref} = 1$. This value for Γ_{ref} is used due to the lack of any previous cost function gradient information to properly evaluate (5.13). Although this represents the most 'relaxed' accuracy level required of the adaptive analysis, i.e. the accuracy requirement is not as stringent as for a value $\Gamma_{ref} < 1$ (which will be the case for the subsequent analysis steps), it is still within acceptable levels.

Not only does the gradient value from the pre-optimization analysis become the initial search direction vector for the line search algorithm, but it is also used for computing the initial value of Γ_{ref}. Using the pre-optimization point as a lower point for the line search interval, the optimizer finds the upper bound point $C^{(bound)}$, $\nabla C^{(bound)}$ adaptively, satisfying (5.12). The details of the line search interval bounding analysis were given in Chapter 2. It must be noted that Γ_{ref} is not updated using the values $C^{(bound)}, \nabla C^{(bound)}$ at any time during the line search algorithm.

For a given design point k in the line search interval, the adaptive analysis computes $C^{(k)}$ and $\nabla C^{(k)}$ with the same accuracy requirements set by the right-hand side of (5.12) as for the bounding point. In fact, the value of the Γ_{ref} parameter is set constant for the entire line search. To strictly adhere to the condition set by (5.13) entails updating the Γ_{ref} parameter for every analysis point on a given line search. This results in a situation where the cubic minimization subroutine will have to deal with design points of varying accuracy on the same line, which may adversely affect the effectiveness of the line search. Hence, Γ_{ref} is updated only at the beginning of a line search using the cost function gradient values from the last valid design point of the preceding line search, as shown in Figure 5.1. This approximation does not reduce the effectiveness of the error calculation, as the changes in the frequency response from one line search to the next are generally not drastic, especially close to the local minimum point.

The resulting CFE produces the cost function and gradient values, which are then checked for the line search termination criteria, either due to geometry re-mesh, or the angle criterion violation by the new $\nabla C^{(k)}$, which were explained in Chapters 3 and 2, respectively. If either condition is satisfied, the current line search is terminated, and the latest ∇C value becomes the search direction for the newly launched line search. The Γ_{ref}

parameter is updated by the optimizer. For the scenario where neither line search termination condition is satisfied, the optimizer finds the new design point $\{g^{(k)}\}$ on the same line and the CFE is done for the same accuracy as the previous design point.

With this implementation, the optimizer exercises control over the solution accuracy through the relative error of the cost function gradient at the adaption frequency, updating Γ_{ref} regularly throughout the optimization. The optimizer also ensures line search efficiency by setting consistent accuracy levels for a given line search, and maintaining cost function continuity by launching new line searches whenever necessary. The CFE module, on the other hand, simply produces the cost function and gradient values to the specifications set by the optimizer, and updates the optimizer on the status of the CFE analysis (pertaining to p-saturation and geometry re-mesh). In the end, the optimizer determines the necessary action for the overall process based on the information sent by the CFE module.

Figure 5.2 shows the steps involved in the CFE analysis associated with each design point set by the optimizer. The steps are similar for all three phases of the line search algorithm, with the exception of the re-mesh scenario, which only occurs for the cubic analysis point. Recall the mesh generation box illustration in Figure 3.3 for details of the re-mesh condition. The FE module is essentially a 'black-box' where it takes the input data provided by the optimizer and produces the cost function and gradient value information with the required accuracy. The CFE module conveys the validity of the adaptive solution to the optimizer through the p-saturation flag: if the adaptive analysis is unable to satisfy the accuracy conditions even with the highest element orders, the solution is rendered invalid, and a flag is set to denote this. The optimizer, in turn, terminates the entire process, as any further improvements to the design are impossible without valid FE solutions. On the other hand, for a good solution, the cost function and gradient values as well as the frequency sweep data are used to compute Γ and handed over to the optimizer. It must be noted the saturation condition is applicable to all three phases of the line search algorithm (as shown in Figure 5.1), which will consequently result in the optimizer terminating the analysis at any stage of the optimization.

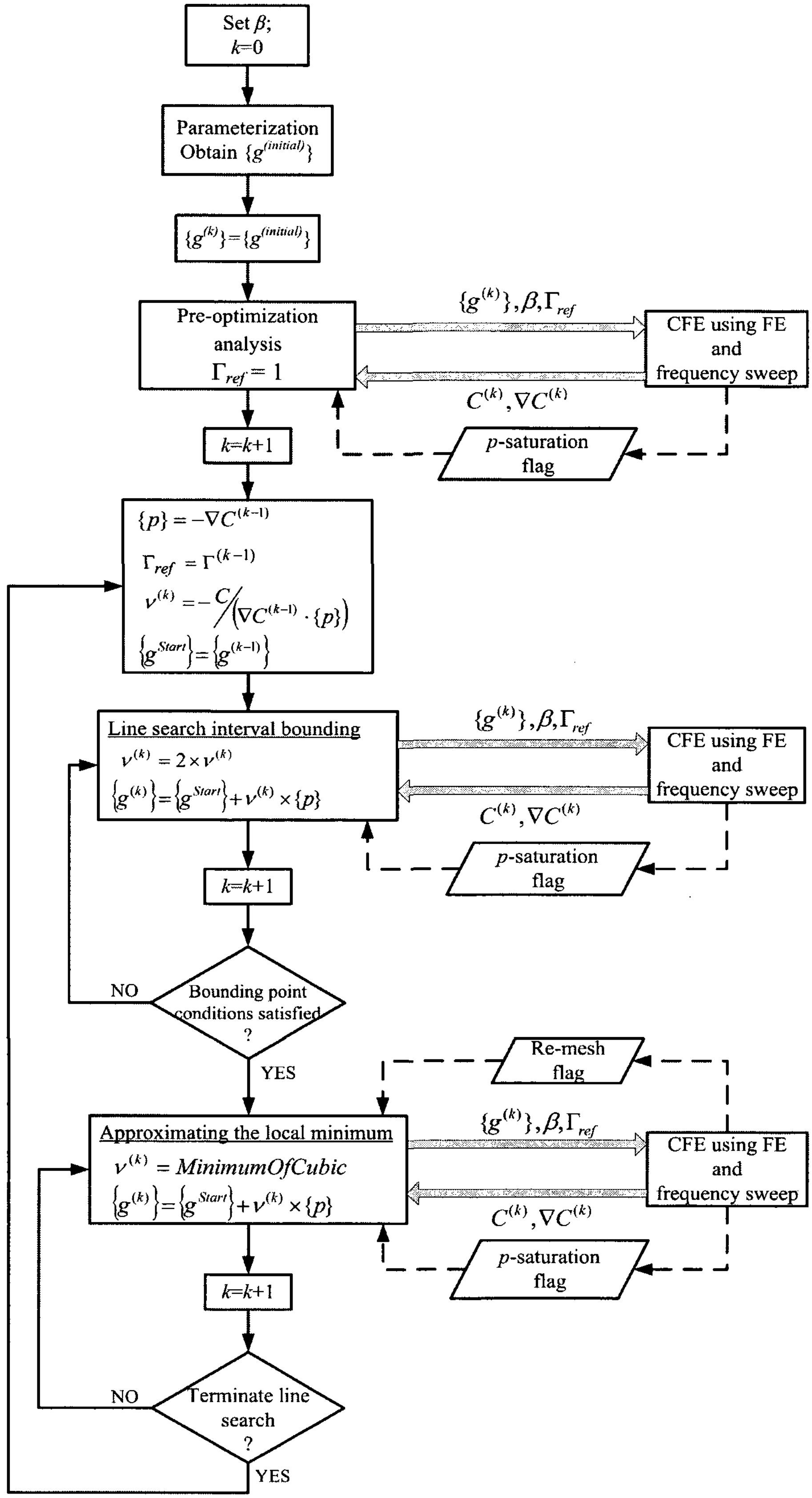

Figure 5.1 Flow diagram for combined adaption-optimization scheme

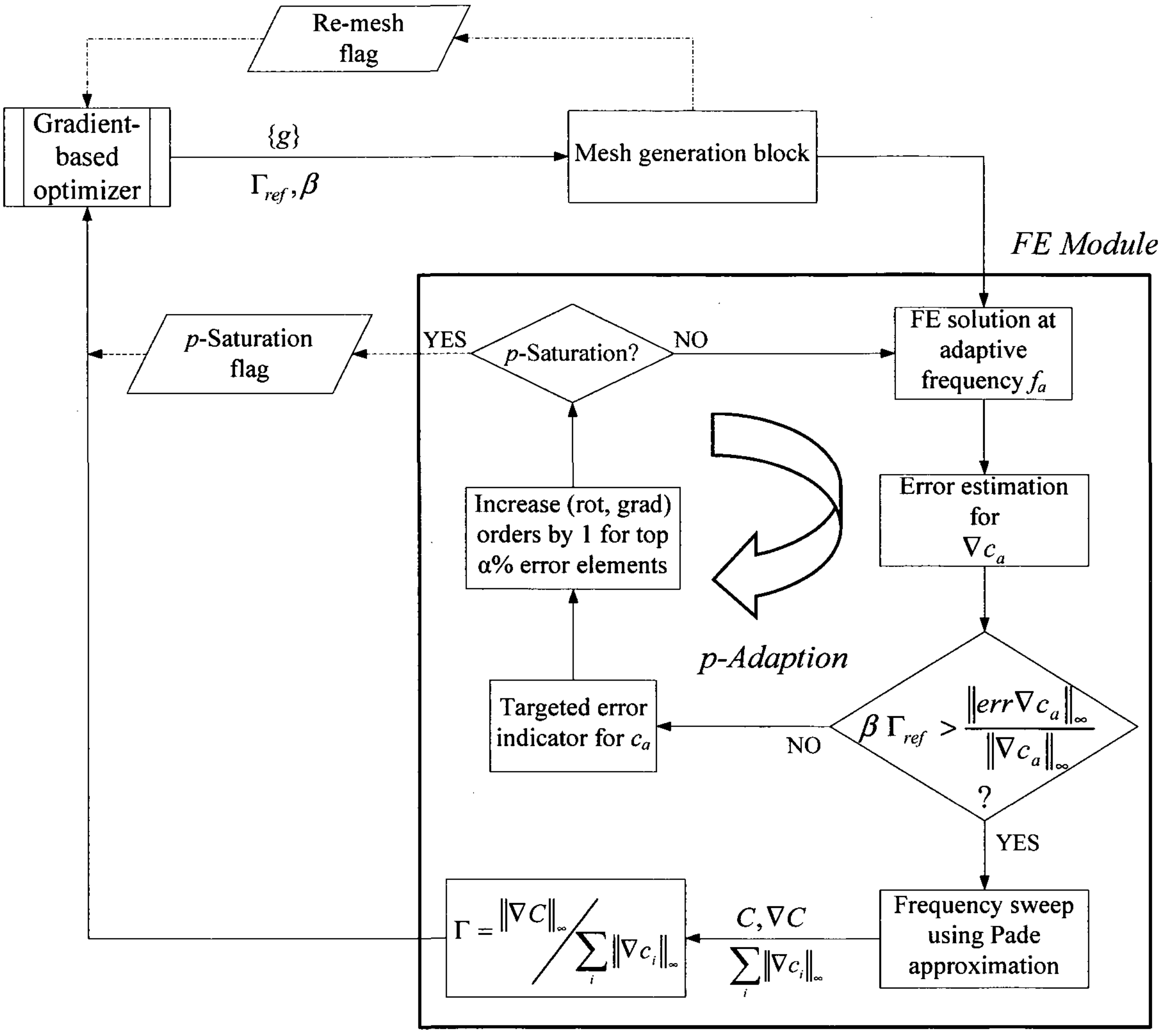

Figure 5.2 The Cost Function Evaluation module for the combined adaption-optimization scheme

5.2 Test cases and numerical results

The adaption-optimization scheme using the accuracy link was tested on a number of 3D microwave problems. The scheme was tested for a single frequency analysis case as well as for frequency-range problems. The device excitation in all cases was the dominant TE_{10} mode for a rectangular waveguide, and the frequency range was selected so that only the dominant mode was allowed to propagate. The cost functions were defined in terms of the scattering parameters in order to minimize the return loss at the ports of the given microwave device. For the single frequency case, the cost function was defined as the square of the magnitude of S_{11} at the adaptive analysis frequency, while for the frequency-range problem it was taken as the average value of the square of the magnitudes of the scattering parameters over 1024 sample points. The optimization was allowed to continue until the adaptive analysis hit p-saturation, i.e., until all elements were at Order 3. To demonstrate the effectiveness of the new scheme, the same design problem was optimized non-adaptively using uniform Order 3 elements throughout, and the computational costs were compared. It must be noted that for the adaptive analysis, the cost function was re-computed using Order 3 elements at each design point, so that the quality of the design could be fairly assessed when comparing to the quality of the design points obtained with the uniform Order 3 optimization.

The figure used to measure the computational cost of an adaptive FE analysis was defined in Section 4.3.2. For the frequency range problems, the cost of the Pade technique must be accounted for as well. The main cost for the Pade frequency sweep is associated with computing the moments of the scattering parameter S_{pq} and the sensitivities dS_{pq}/dg at the adaptive frequency. How these moments are obtained from the derivatives of the field $\{E^{(q)}\}$ is explained in Section 3.3.1, and the references mentioned therein. Calculating the derivatives involves the solution of a matrix system of similar size as the last adaptive step for the given design point. This solution is repeated M times, where M is the number of moments. Hence, the total cost for one design point will be the sum of a) the cost of the adaptive FE analysis, and b) M x the cost of solving the last adaptive step. Here, a value of M=10 was used. The cumulative cost of the overall analysis would then be the sum of the costs for all design points (pre-optimization, bounding and cubic points) during the analysis.

As mentioned before, a value of β=0.2 (which represents a relative error in $\nabla C^{(k)}$ of 20% in (5.1)) was used for all the examples. As with the test cases in Chapter 3, the threshold angle was set to θ_{th}=45°, which was deemed a significant enough deviation between $\nabla C^{(k)}$ and the search direction $\{p^{(k)}\}$ to terminate the current line search and initiate a new one.

5.2.1 Single frequency analysis of a waveguide impedance transformer

The return loss of the waveguide impedance transformer (shown in Figure 3.10) is to be minimized. The design variables of interest and their initial values are indicated in the figure. The algorithm was applied to a single frequency problem, which for this case happens to be the adaptive analysis frequency. For this scenario, the cost function and gradient can be defined as follows:

$$C = c_a$$

$$\nabla C = \nabla c_a \tag{5.14}$$

where the subscript a indicates the adaptive analysis frequency point f_a. Since the analysis is at a single frequency point, (5.13) is evaluated using the value at a single point, which translates into Γ = 1 in (5.11), with both the numerator and denominator being equal to each other. Thus, the accuracy requirement, (5.12), is set by the parameter β alone.

The cost function reduction for the transformer is plotted in Figure 5.3 as a function of the cumulative cost. The result shows a significant decrease in the cost function value in the early stages of the optimization (the first four or five steps), even though the element orders used in the adaptive analysis are low. Yet it manages to outperform the Order 3 results in terms of computational costs to attain similar improvements in the design performance. Beyond this point, the adaptive analysis takes more steps than the fixed order method to reduce the cost function.

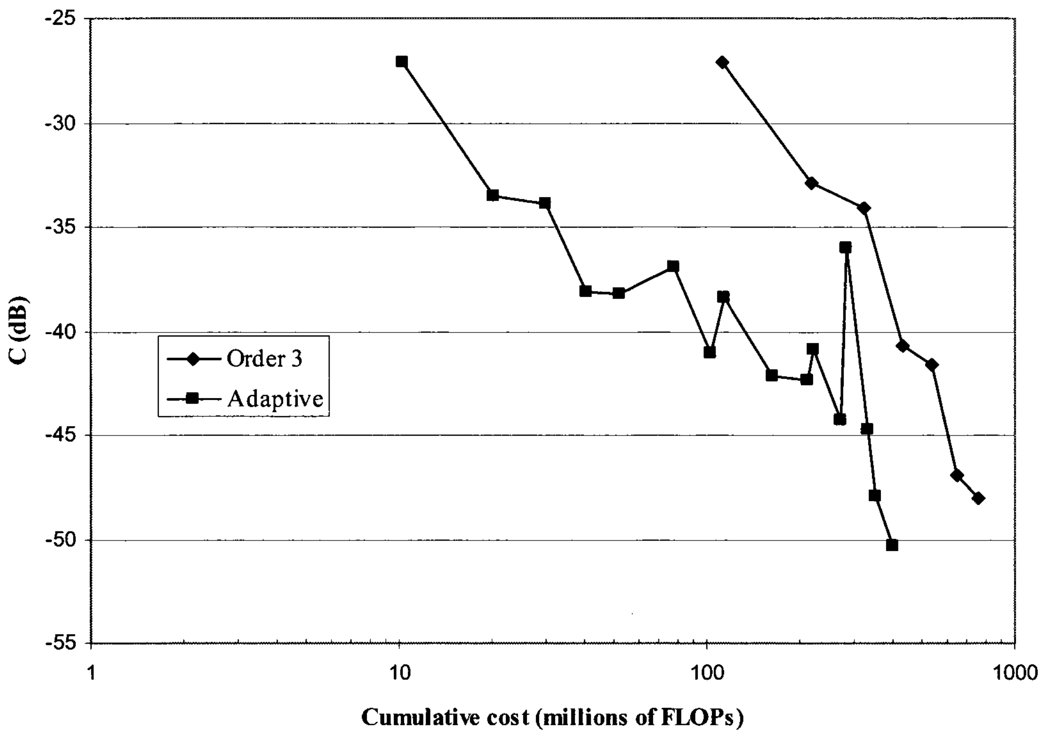

Figure 5.3 Cost function reduction for the transformer in Fig. 3.10 using both the Order 3 optimization and the adaptive method for the single frequency analysis. Initial $\{g\}^T = \{22, 6, 10.44\}$ mm. Analysis frequency = 10GHz. Initial number of tetrahedra=1700.

Table 5.1 gives the final design parameters produced by both optimization schemes, as well as the cumulative cost for each method. The similarity in both sets of design parameter values indicates the optimization was successful in guiding the design to a local optimum.

Table 5.1 Final design parameter values (in mm), cost function values (in dB) and cumulative cost (in millions of FLOPs) for Order 3 and adaption-optimization schemes applied on the waveguide transformer (single frequency)

	Final	
	Order 3	Adaptive
$\{g\}^T$	$\{21.81,5.72,10.10\}^T$	$\{21.78,5.72,10.19\}^T$
C	-48.0	-50.4
Cumulative cost	763	401

The final cost function value for the adaptive scheme is also similar to that of the Order 3 analysis, with a computational cost that is almost half of that for the high order analysis, thus validating the idea that the accuracy link leads to computational cost savings.

The single frequency results were reported in [108].

5.2.2 E-plane bend analysis at two frequency points

The next example is a 2D optimization problem using the E-plane bend shown in Figure 3.4. The analysis of this device was conducted at two frequency points, selected at the two ends of the operating frequency range. The cost function is defined as a combination of the S_{11} values at the two ends of a given frequency range. Defining the cost function in such a manner allows the optimizer to adjust the frequency response by balancing the end point values while trying to minimize them.

The Pade method is applied to approximate the frequency response using 1024 sample points (i = 0 -> 1023) spread evenly across the range, and the two end points are selected to define the cost function and the gradient:

$$\begin{aligned} C &= c_1 + c_2 \\ \nabla C &= \nabla c_1 + \nabla c_2 \end{aligned} \tag{5.15}$$

where f_1 = 8.25 GHz and f_2 = 13.5 GHz correspond to the frequency values at the two end points, i = 0 and i = 1023, respectively, of the Pade approximation of the range. The adaptive frequency is selected at the center point of the range at f_a = 10.875 GHz, and the analysis is required to satisfy the accuracy condition of (5.12). Evaluating (5.13) now involves the values of the cost function gradient at f_1 and f_2. The reduction in the cost function is shown in Figure 5.4.

From the graph, it can be seen that although the initial point for both curves from the pre-optimization step has the same return loss, the Order 3 analysis is costlier than the adaptive one. The adaptive analysis is able to achieve significant cost function reductions with the next two points, each representing a line search. The last three analysis points of the curve belong to the final line search, where the analysis appears to be slowing down.

The final design point for both methods attain return loss values around -31.5 dB, with the Order 3 analysis requiring three times the computational cost of the adaptive analysis.

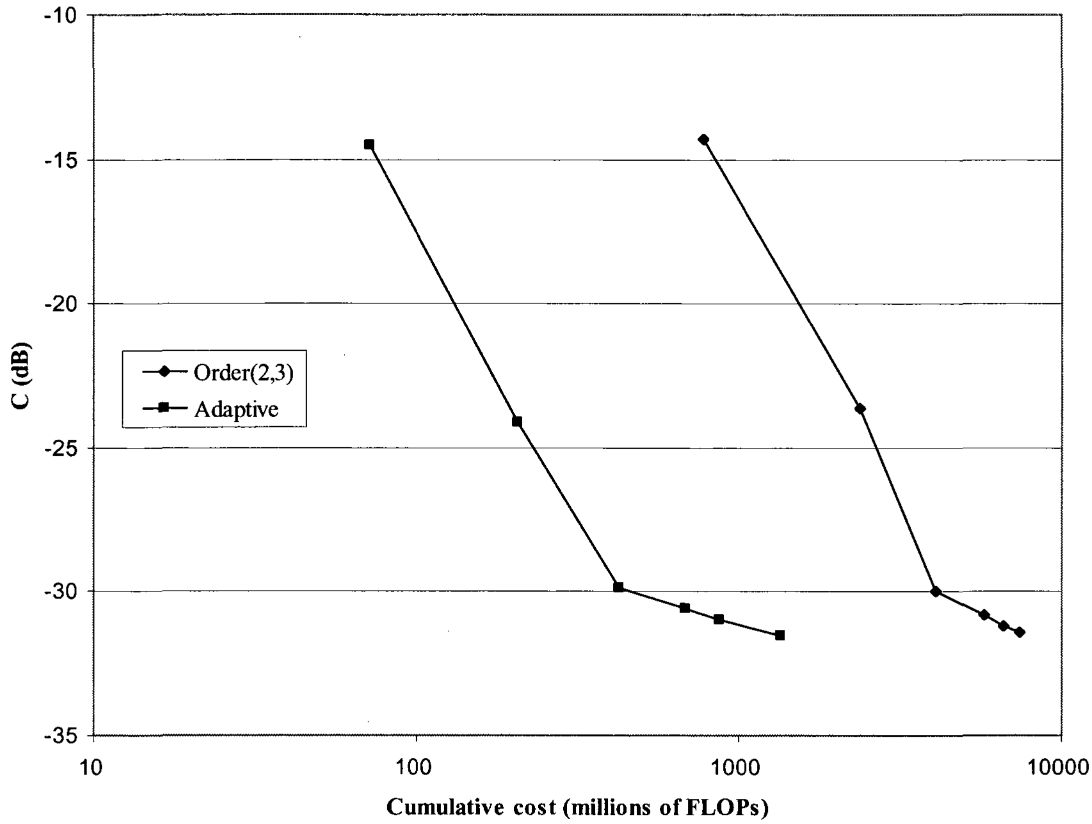

Figure 5.4 Cost function reduction for the E-plane bend for both adaptive and uniform Order 3 analysis. Initial $\{g\}^T = \{1, 1\}$ mm. Adaptive frequencies: f_a=10.875 GHz. Initial no. of tetrahedra=1400

The cost function gradient values achieve a reduction of two orders of magnitude during the optimization analysis, with initial values {0.1292, 0.1289}, and final values {0.0037, 0.0040}. This reduction in the gradient values is reflected in the higher accuracy demands imposed on the analysis as the optimization progresses.

Figure 5.5 shows the frequency response of the device over several optimization steps. Curve 1 represents the frequency response obtained for the initial design point; curve 2 an intermediate point during the optimization (at the end of the second line search), and curve 3 the final optimized design configuration. The optimizer seems to have adjusted the design so that the resonance dip in the frequency response has been positioned to reduce the overall cost function value. However, the final curve appears unbalanced at the end points, contrary to what was expected from the optimization. The reason could be that the analysis hit saturation (thus, termination of the optimization) before achieving this balance.

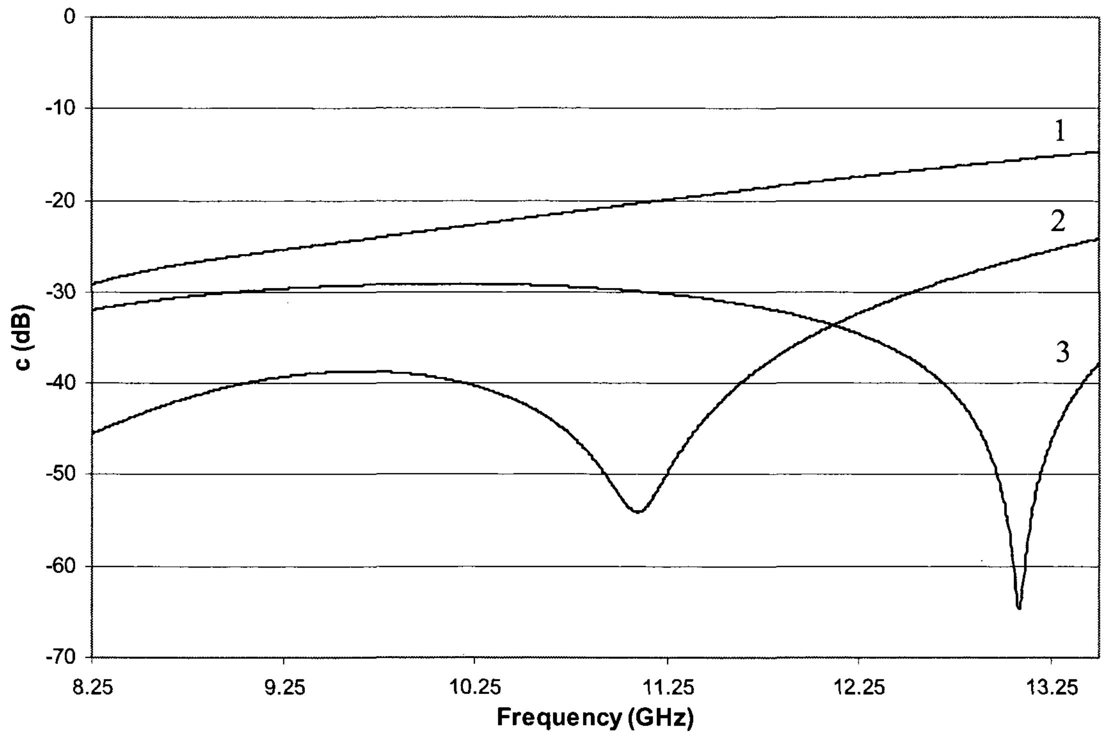

Figure 5.5 Frequency response of the E-plane bend at various design points during the combined adaption-optimization analysis; the cost function value at each sample point, $c_i=|S_{11}|^2$

Table 5.2 gives the final design parameter and cost function values for both the adaption-optimization and uniform Order 3 optimization, along with their respective cumulative computational costs. Once again, the cost savings achieved with the adaptive analysis while obtaining similar final design points validates the use of the accuracy link.

Table 5.2 Final design parameters (in mm), cost function values (in dB) and cumulative cost (in millions of FLOPs) from Order 3 and adaption-optimization analysis of the E-Plane bend problem

	Final	
	Order 3	Adaptive
$\{g\}^T$	$\{2.1829, 2.1826\}^T$	$\{2.1946, 2.1891\}^T$
C	-31.41	-31.51
Cumulative cost	7,421	1,356

5.2.3 Waveguide transformer problem over a frequency range

The combined adaption-optimization analysis is applied to the waveguide transformer from Section 5.2.1 to improve the device performance over a frequency range. Preliminary results were reported in [109].

For this example, the expression in (5.12) was used for accuracy control, and the Γ-parameter was approximated using (5.13). For the pre-optimization point, Γ=1, as shown in Figure 5.1. The initial design parameter values and the adaptive frequency were the same as those presented in Section 5.2.1. The cost function and gradient definitions are given by (5.2) and (5.3). Figure 5.6 shows the cost function reduction over the course of the optimization analysis.

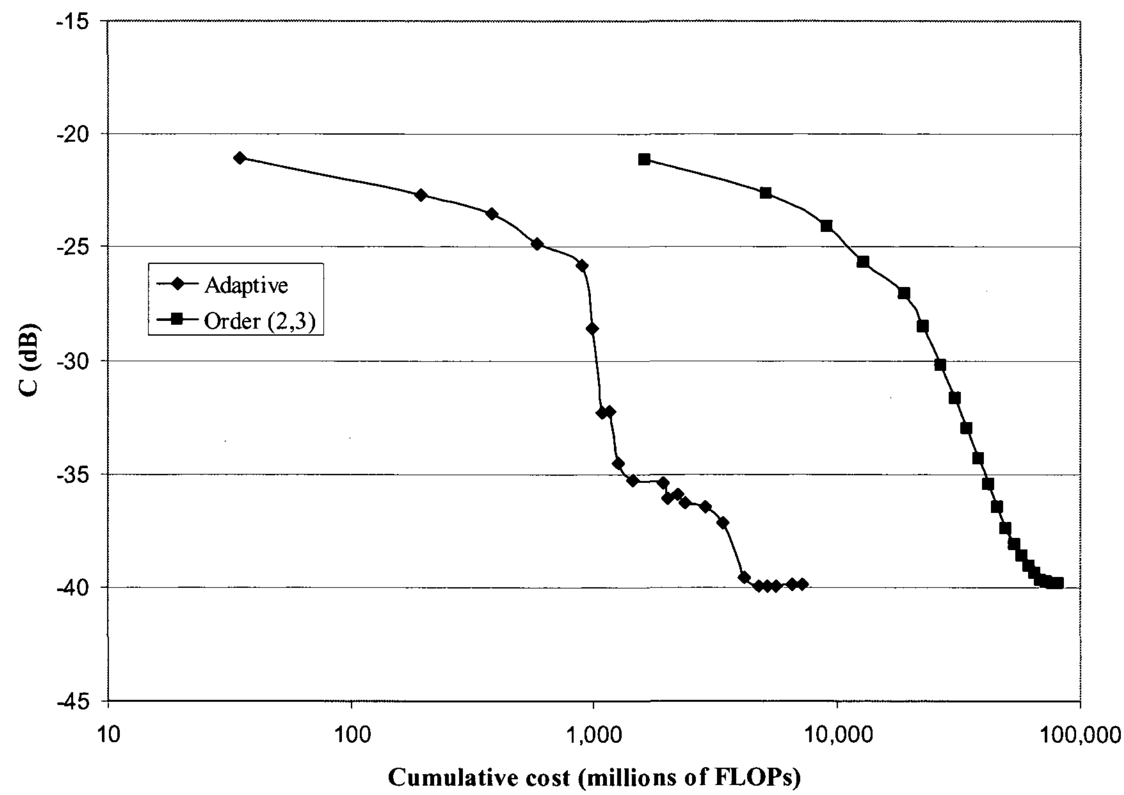

Figure 5.6 Cost function reduction for the waveguide transformer using both the adaptive and uniform Order 3 analysis. Initial $\{g\}^T$= {22, 5.94, 7.83} mm. Adaptive analysis frequency=10.25 GHz. Initial number of tetrahedral=2200

The adaptive scheme required 16 line searches before being terminated due to saturation. Both methods attain a 20 dB reduction in the cost function values, but with the Order 3 analysis being significantly (approximately 17 times) more expensive. From Figure 5.6, the Order 3 curve shows a smooth reduction over the analysis steps, while the adaptive results are more uneven. This may be attributed to the cost function gradient values calculated at low accuracies in the early stages of the optimization using lower order elements, and hence less accurate than the Order 3 analysis values. However, they do satisfy the accuracy levels set by the optimizer at that stage, and help direct the line search in the proper direction in design space. Even more, these early adaptive stages achieve the significant reduction in the cost function for a much lower cost. The link assures an increase in the solution accuracy in the late stages of the optimization with a decreasing gradient: the gradient vector is reduced by two orders of magnitude, from an initial value $\{5.817, 9.865, 6.949\}$ to the final value of $\{0.0340, 0.0201, 0.0258\}$.

Figure 5.7 illustrates the changes in the frequency response of the device throughout the adaption-optimization. For clarity, only a few response curves are shown. Curve 1 represents the response of the initial design, obtained with the pre-optimization analysis, and curve 4 is the response of the final design. Curve 2 and 3 represent the design points obtained at the end of line searches 7 and 9, respectively. The optimizer has once again adjusted the final design so that the resonance dip is positioned in the middle of the frequency range, and the two end points balanced.

Table 5.3 shows the final values for the design parameter values and the cost function values after both the adaptive and uniform order optimization. The significant computational savings and the close correspondence between the design point values validate the effectiveness of the optimization scheme incorporating the accuracy link.

Table 5.3 Final design parameter values (in mm), cost function values (in dB) and cumulative computational cost (in millions of FLOPs) from Order 3 and adaption-optimization analysis of the waveguide transformer (frequency range)

	Final	
	Order 3	Adaptive
$\{g\}^T$	{23.06,5.95,9.3}	{23.08,5.95,9.32}
C	-39.84	-39.91
Cumulative cost	81,260	7,233

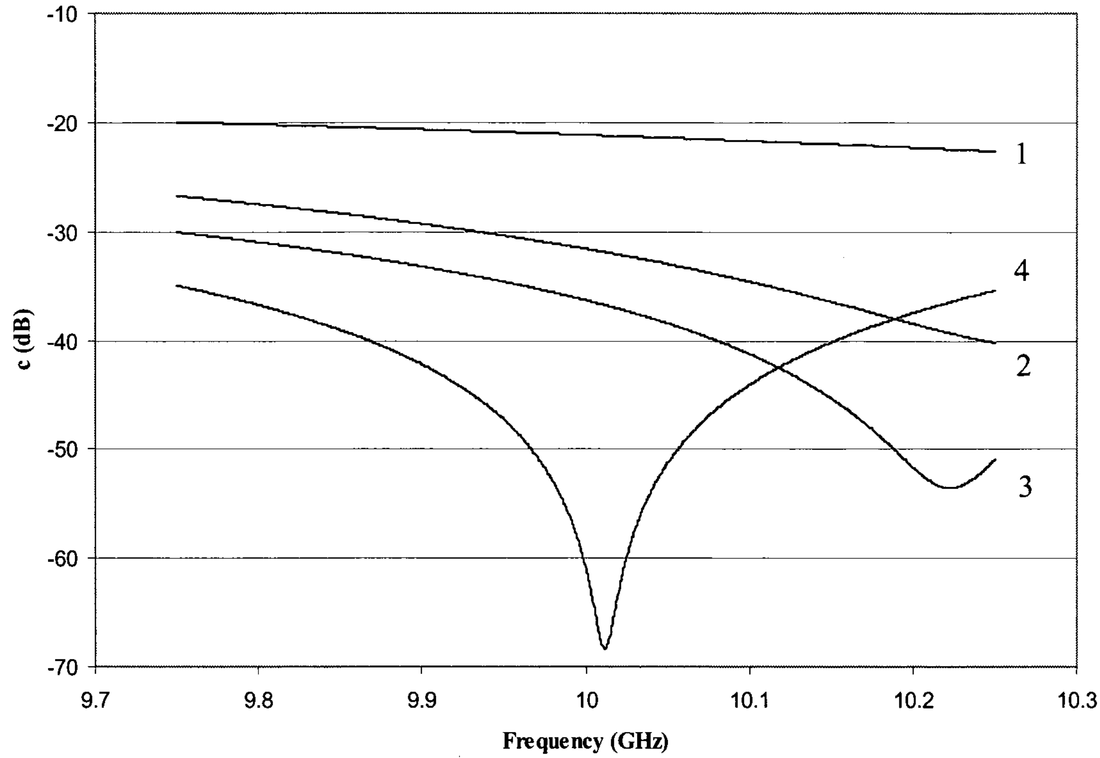

Figure 5.7 Frequency response curves of the waveguide transformer during the adaption-optimization analysis; the curves are identified with their line search number

5.2.4 Magic-T optimization over a frequency range

The final test case for the new adaption-optimization scheme is the magic-T problem shown in Figure 4.13. The cost function definition is based on the return loss at two ports: ports 1 and 4. Due to the symmetries of the magic-T, a reduction in the return loss at port 1 and port 4 combined will result in a reduction in the return loss at all the ports of the device [38], [110]. Hence a cost function definition that is a combination of both parameters will properly characterize the given device, and minimizing it will improve the overall performance. The design parameters used for the optimization are the same as was defined for the problem in Chapter 4. An operating frequency range of 7 GHz to 11 GHz was used, and the adaptive analysis frequency selected as 9 GHz. Equations (5.12) and (5.13) were used for accuracy control. Results for this problem were also presented in [109]. Figure 5.8 shows the reduction in the cost function during the optimization.

The adaption-optimization scheme ran for only four line search steps before the high accuracy requirements led to p-saturation. The increasing accuracy demands are reflected in the increase in the cumulative cost over the course of the analysis run: by the time the analysis reaches the final adaptive point, the cumulative cost comes closer to matching the Order 3 analysis cost. Even so, the adaptive method reduces the cost function at almost half the cost of the Order 3 analysis. The cost function could be reduced further with the help of a much finer discretization. The presence of numerous re-entrant corners throughout the structure, and the difficulty in trying to approximate a circular post with a relatively coarse mesh may have contributed to high discretization errors preventing the successful analysis of the problem.

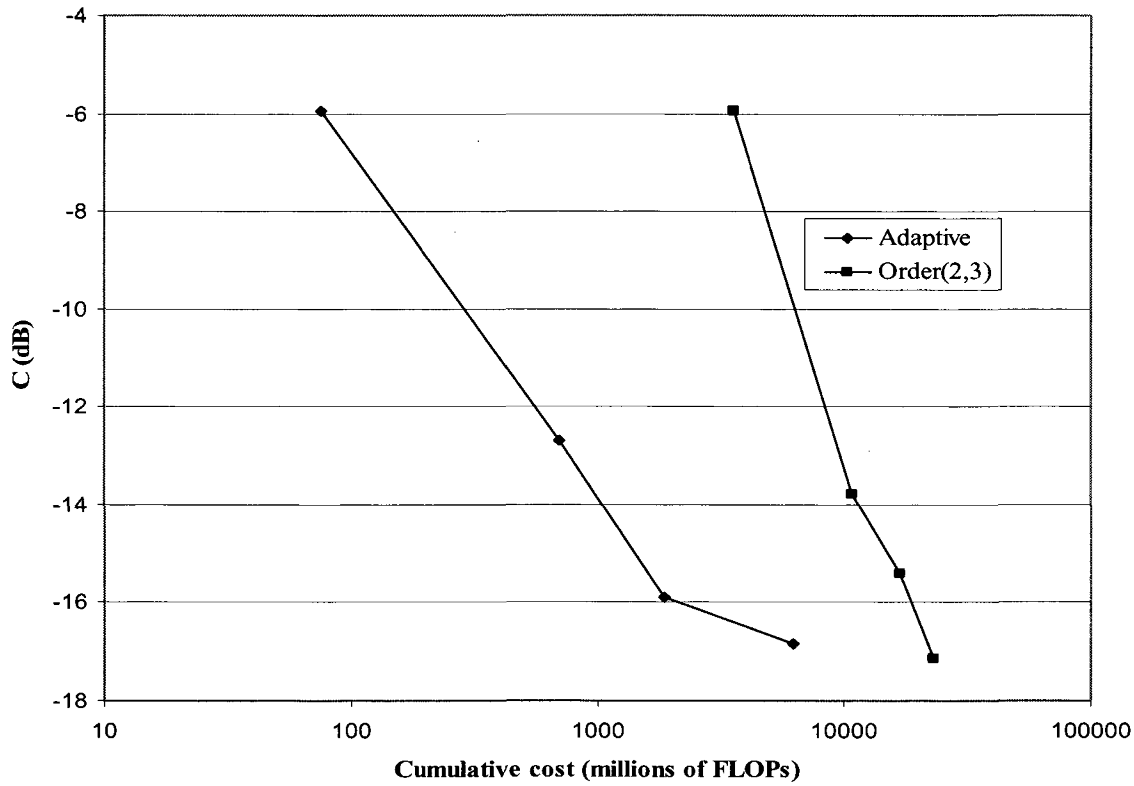

Figure 5.8 Cost function reduction for the magic-T shown in Figure 4.13 using both the adaptive analysis and the uniform Order 3 elements. Initial $\{g\}^T$ = {8.225, 12.7, 3.0, 3.556, 8.382, 1.60} mm. Adaptive frequency = 9GHz, Initial number of tetrahedra=6173

The plot in Figure 5.9 shows the change in the frequency response of the magic-T after the optimization analysis. Although the reduction is not as significant as in previous examples, the changes are based on design parameter values after only four steps.

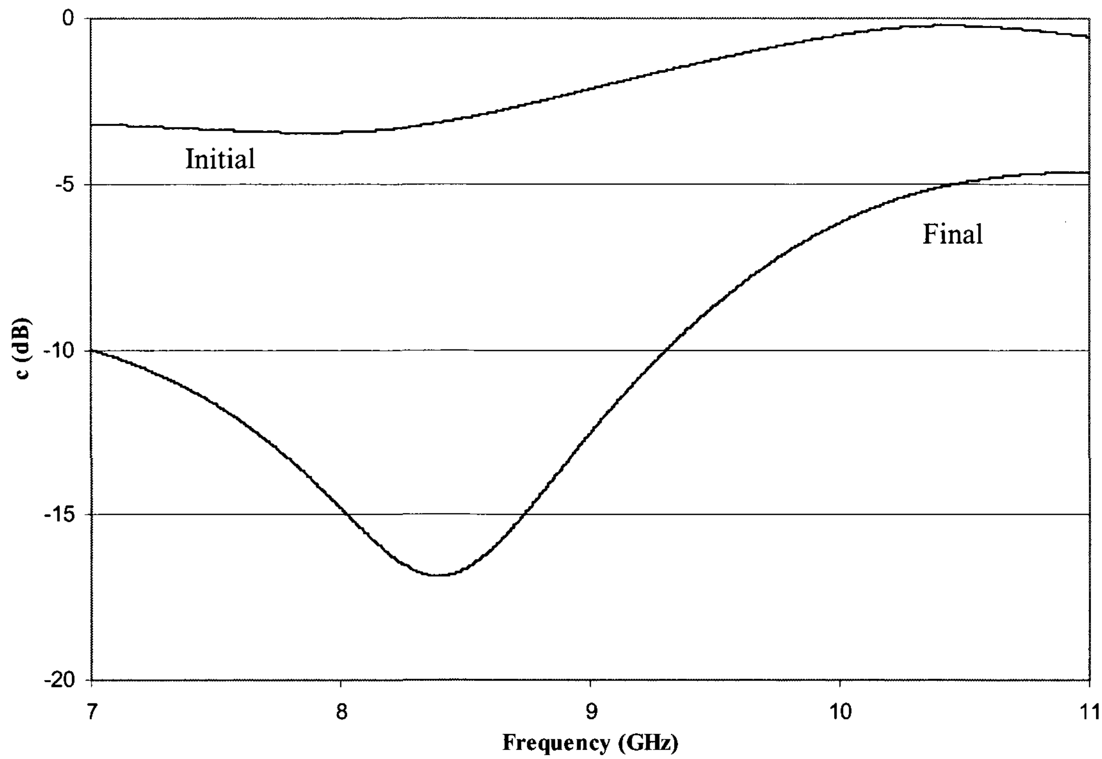

Figure 5.9 Frequency response for magic-T

Table 5.4 gives the final values of the cost function as well as the design parameters after optimization using the adaptive technique and the Order 3 analysis.

Table 5.4 Final design parameter values (in mm), cost function values (in dB) and cumulative computational cost (in millions of FLOPs) from Order 3 and adaption-optimization analysis of the magic-T over a frequency range.

	Final	
	Order 3	Adaptive
$\{g\}^T$	{6.96,12.48,1.35,3.86,6.37,0.53}	{7.46,12.58,1.50,3.45,7.06,0.89}
C	-17.14	-16.85
Cumulative cost	23,037	6,154

5.3 Discussion

The results in this chapter demonstrate a combined adaption-optimization scheme with an accuracy link that exerts control over the solution process. All examples were able to attain significant cost function reductions over a small number of steps. In most cases, the cost gradients were reduced considerably, up to two orders of magnitude from the initial set. This resulted in higher accuracy demands in the late stages, and eventually *p*-saturation. The E-plane problem in Section 5.2.2 is an example of the gradient being reduced rather quickly leading to early termination. The issue of high discretization error in complex 3D structures may also play a role in the early termination of the CFE, e.g., the magic-T problem. However, it is this apparent 'drawback' that is one of the benefits of the accuracy link formulation: preventing numerically invalid solutions from contaminating the design optimization process. As soon as the CFE solutions are corrupted with 'numerical noise' when trying to meet exorbitant accuracy demands, or by discretization error due to severe under-refinement, the optimizer terminates the analysis.

One weakness of the link is its inability to deal with the minimax type of optimization problems. The examples in Chapter 3 are of this type, however without the accuracy link. Trying to optimize this type of cost function with the accuracy link can lead to the breakdown of the scheme, due to the absence of any reduction in the cost function gradient. An example is shown in Figure 5.10, where the cost function is essentially v-shaped, with the local minimum falling at the intersection point.

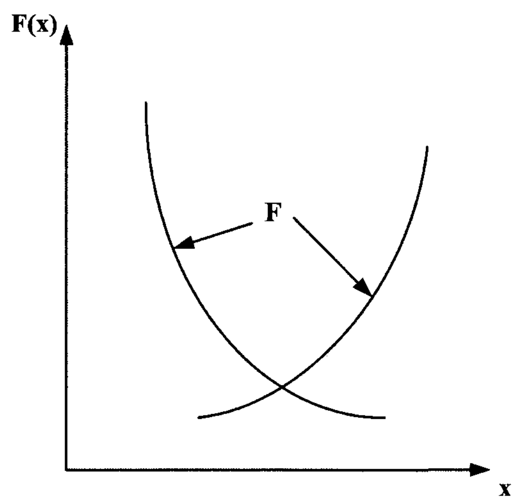

Figure 5.10 An example of a minimax type of optimzation problem

Although the steepest-descent optimizer will converge to the local minimum (as was demonstrated with the results in Chapter 3), the accuracy link will fail since the gradient value may not have been reduced significantly, if at all, and this will not have any impact on the accuracy requirements of the analysis.

6. Conclusions

This thesis is an attempt to formulate a computational scheme for the design optimization of 3D microwave problems. This work illustrates the computational efficiency with which the performance characteristics of a given 3D device can be improved provided certain control is exercised over the accuracy requirement at each analysis step. The underlying idea is that the required accuracy levels for the numerical solution need not be constant throughout the entire process, but rather, should be determined by the requirements of the optimization at every stage, which translates into computational savings. The two main components that form the accuracy-controlled optimization system in this research are: a steepest-descent method using a cubic minimization line search technique, and a p-adaptive FE analysis method to compute the cost function and its gradient. The accuracy control for the numerical analysis is achieved by linking the FE module back to the optimizer.

The optimizer is custom-made, with special provisions made for enhancing the efficiency of the overall analysis. One such provision is allowing the line search to terminate provided certain conditions are met, rather than performing an exact line search, as explained in Chapter 2. This avoids extra line search steps, especially in the early stages, resulting in cumulative cost savings. Another modification is related to maintaining the cost function continuity during the analysis: when the mesh topology changes from one optimization step to the next, the optimizer terminates the current line search and initiates a new one. This resulted in the mesh topology being fixed during a given line search in order to keep the cost function as smooth as possible. Cost function discontinuities can undermine the line search algorithm and even cause it to fail. The relatively small changes in the geometry were dealt without altering the topology through a mesh perturbation process, as explained in Chapter 3.

Calculating the cost function gradients requires the sensitivity information of the mesh vertices that are associated with each design parameter. A custom implementation using VBScript, the ElecNet mesh generator and C++ code was used for this purpose. The vertex sensitivity values also play an important role in the mesh perturbation process to help determine the amount by which the nodes should be moved to reflect the changes in the geometry, as explained in Chapter 3.

The FE analysis module is called by the optimizer at various stages to compute the cost function and its gradients. A modified version of the Adjoint Variable Method using FE is used to calculate the gradients with no additional solutions, regardless of the number of design parameters. The Pade approximation technique is used to obtain the frequency response of the device at a design point, from which the cost function and the gradients are computed. An optimization implementation employing these gradients proved to be quite effective in finding a local minimum, as shown in Chapter 3.

An error indicator and an error estimator are key components for the *p*-adaptive FE analysis. An indicator which makes use of the properties of hierarchal elements to target the error in the scattering parameters in each element was developed. The indicator identifies those elements with high error at each adaptive step in order to further refine them and increase solution accuracy. An accurate error estimator gauges the accuracy of the overall solution and helps to determine the proper termination point for the adaptive analysis. Since the accuracy of the gradient is important for the efficiency of the optimization, an error estimator is developed to assess the quality of the gradient values computed by the FE solver. The results in Chapter 4 demonstrate how effective the new indicator scheme is in reducing the error in the scattering parameters, and how the new *a-posteriori* estimator is able to produce accurate error estimates for the gradient.

The accuracy link is described and demonstrated in Chapter 5. The formulation allowed for cost functions based on a frequency sweep, but is general enough to be applied to single frequency problems as well. The implementation involved assembling the components developed in Chapters 3 and 4 to form the combined adaption-optimization system. Essentially, the link allows the optimizer to keep a tight check on the computational effort by dictating the accuracy levels for the adaptive FE solution at each optimization step so as to maintain the efficiency of the overall analysis. The link also allows the optimizer to terminate a line search, or even the entire optimization analysis, based on the status of the FE solver and mesh. The benefit of this is that the optimizer can prevent solutions from being corrupted by numerical noise, as explained in Chapter 5.

A number of test cases using 3D microwave devices are used to gauge the efficiency of the new link formulation. The results obtained with the combined adaption-optimization scheme are compared to results obtained with the optimization analysis using uniform high order, non-adaptive FE solutions. The return loss is used to define the cost function for all the problems. In all cases the new combined scheme is able to achieve significant computational cost savings compared to the non-adaptive high order scheme. An interesting observation for all test cases is that the optimizer is able to attain the same level of cost function reduction as the uniform high-order analysis, even in the early stages of the optimization when the accuracy requirements for the FE solution are low, but at a much smaller cumulative cost. As a best-case example, for the impedance transformer example in Chapter 5, by the midway point of the analysis, both schemes reduce the cost function by approximately 14 dB from the initial point, but with the adaptive scheme achieving it at a much lower cost, with computational savings of a factor of 25. Although the other examples do not show similar cost savings, they are still significant enough to validate the efficiency of the adaptive scheme.

6.1 Original contributions

The following are the original contributions of this thesis:

1. The development of a direct optimization scheme for cost functions depending on device performance over a frequency range [52]. The scheme makes use of gradients efficiently calculated over a frequency range using FE analysis and Pade approximation techniques.
2. An error indicator that uses the properties of hierarchal elements to target the S-parameter error in every element of the FE mesh for the p-adaptive analysis [85].
3. An error estimator that accurately approximates the error in the design sensitivities that are computed by the p-adaptive FE analysis [98].
4. The implementation of a new accuracy link between an optimizer and the p-adaptive FE analysis module [108], [109] for the cost efficient design optimization of 3D microwave problems, incorporating points 1-3.

6.2 Future work

Various improvements can be made to the accuracy link implementation done for this thesis that might be of interest for future research.

Application of the accuracy link to problems that involve h-adaptive techniques is an interesting avenue of research to pursue. For this thesis, the p-adaptive technique is used due to the availability of hierarchal elements, and the error indicator and estimator formulations took advantage of the properties of these elements. But the theory is easily applicable to h-adaptive problems, with an indicator and estimator tailored to suit h-adaption. A combination of the h- and p-type adaptive schemes linked with the optimizer will make an even more powerful optimization scheme. An interesting question here is how to split the task of meeting accuracy demands at each step between the two adaptive schemes; for example, finding the best mesh refinement strategy to meet the accuracy requirements in the early stages as opposed to the late stages of the optimization.

The ability to handle minimax type of optimization problems, which has considerable applications in many practical situations, is also desirable. The major issue with this class of problems is the discontinuous cost functions and the inability of first order methods to deal with them. Although several methods have been proposed to tackle this issue and converge to a local solution (see [111], [112], and the references within), from an accuracy link perspective, the issue is that the gradient, in general, does not equal to zero at the local minimum, thus undermining the idea behind the accuracy link proposed for this thesis. Thus finding a link formulation to overcome this short-coming will be a possible research topic of considerable interest.

Combining the accuracy link with stochastic optimization schemes could improve the computational efficiency of the process of finding a global minimum. Using a combination of gradient based techniques and stochastic methods could lead to an optimization scheme that is far less expensive than the stochastic algorithms while being able to find the global minimum, a property that is lacking in gradient-based methods. For maximum efficiency, how the analysis can be split between the two techniques at various stages of the optimization is a research topic on its own. Formulating a "hybrid" link that uses information from both the deterministic and stochastic optimizers to set the accuracy for the numerical analysis will also become necessary.

Modifying the scheme to include open structure problems, such as scattering problems and antenna radiation problems is yet another possibility for further research. Using a Perfectly Matched Layer (PML) [113] or Absorbing Boundary Conditions (ABC) [114] to truncate the computational domain, then applying the combined adaption-optimization scheme will lead to interesting results. Recent work in [115] demonstrated an FE-based approach for the optimization of planar microwave circuits using first-order ABCs to truncate the problem domain. However for such problems, the cost functions are better defined in terms of the radar cross section, and as such, would require new error estimators and indicators to implement the link. This is an issue that is relevant when applying the adaption-optimization link to problems in any other area of engineering.

References

1. S.S. Rao, *Engineering Optimization, Theory and Practice*, 3rd edition, John Wiley and Sons, 1996.
2. G. Matthaei, L. Young and E. M. T. Jones, *Microwave Filters, Impedance-Matching Networks and Coupling Structures*, McGraw-Hill Book Company, MA, 1985.
3. D.A. Lowther, "Knowledge based and numerical optimization techniques for the design of electromagnetic devices," International Journal of Numerical Modeling: Electronic networks, devices and fields, Vol.9, pp. 35-44, 1996.
4. F. Arndt, J. Brandt, V. Catina. J. Ritter, I. Rullhusen, J. Dauelsberg, U. Hilgefort, and W. Wessel, "Fast CAD and optimization of waveguide components and aperture antennas by Hybrid MM/FE/MoM/FD methods – State of the art and recent advances", IEEE Transactions on Microwave Theory and Techniques, Vol.52, No.1, pp. 292-305, 2004.
5. W.C. Chew, J.-M. Jin, E. Michielssen and J. Song, *Fast and Efficient Algorithms in Computational Electromagnetics*, Boston, MA: Artech House, 2001.
6. H.-B. Lee and T. Itoh, "A systematic optimum design of waveguide-to-microstrip transition", IEEE Transactions on Microwave Theory and Technique, Vol.45, No.5, pp. 803-809, 1997.
7. E.J. Haug, K.K. Choi and V. Komkov, *Design Sensitivity analysis of Structural Systems*, Orlando, FL: Academic, 1986.
8. A.D. Belegundu and T.R. Chandrupatla, *Optimization concepts and applications in Engineering*, Upper Saddle River, NJ: Prentice Hall, 1999.
9. H. Akel and J.P. Webb, "Design sensitivities for scattering-matrix calculations with tetrahedral edge elements", IEEE Transactions on Magnetics, Vol.36, No.4, pp. 1043-1046, 2000.
10. M.M. Gavrilovic and J.P. Webb, "Accuracy control in the optimization of microwave devices by finite element methods", IEEE Transactions on Microwave Theory and Techniques, Vol.50, No. 8, pp. 1901-1911, 2002.

11. J.W. Bandler, R.M. Biernacki, S.H. Chen, L.W. Hendrick, and D. Omeragic, "Electromagnetic optimization of 3-D structures", IEEE Transactions on Micorwave Theory and Techniques, Vol.45, No.5, pp. 770-779, 1997.
12. J.W. Bandler, "Optimization methods for computer-aided design," IEEE Transactions on Microwave Theory and Techniques, Vol.17, pp. 533-552, 1969.
13. V. Rizzoli, A. Costanzo, D.Masotti, A.Lipparini and F. Mastri, "Computer-aided optimization of non-linear microwave circuits with the aid of electromagnetic simulation," IEEE Transactions on of Microwave Theory and Techniques, vol.52, no.1, pp. 362-377, 2004
14. Y. Kuwahara, "Multi-objective optimization design of Yagi-Uda antenna," IEEE Transactions on Antennas and Propagation, Vol.53, No.6, pp. 1984-1992, June 2005.
15. Ansoft Corporation, "Parametrics and optimization using Ansoft HFSS," Microwave Journal, Vol. 42, No. 11, Horizon House Publications Inc., pp.156-166, November 1999.
16. P. Di Barba, B. Forghani, and D.A. Lowther, "Discrete-valued design optimization of a multiple-coil inductor for uniform surface heating," International Journal for Computation and Mathematics in Electrical and Electronic Engineering, Vol. 24, No. 1, pp. 271-280, 2005.
17. K. Preis, C. Magele, and O. Biro, "FEM and Evolution Strategies in the optimal design of electromagnetic devices," IEEE Transactions on Magnetics, Vol.26, No.5, pp. 2181-2183, September 1990.
18. J.H. Holland, *Adaptation in Natural and Artificial Systems*, Ann Arbor: University of Michigan Press, 1975.
19. Y. Rahmat-Samii and Eric Michielssen, Editors, *Electromagnetic Optimization by Genetic Algorithms*, Toronto: John Wiley and Sons, 1999.
20. S. Kirkpatrick, C.D. Gelatt Jr., and M.P. Vecchi, "Optimization by Simulated Annealing," Science Vol.220, pp. 671-680, 1983.
21. Q.J. Zhang and K.C. Gupta, *Neural Networks for RF and Microwave Design*, Boston: Artech House, 2000.

22. J. Robinson and Y. Rahmat-Samii, "Particle Swarm Optimization in electromagnetics," IEEETransactions on Antennas and Propagation, Vol.52, No.2, pp. 397-407, 2004.

23. Wei-Chung Weng, F.Yang and A. Elsherbeni, *Electromagnetics and Antenna Optimization using Taguchi's Method*, Morgan & Claypool, 2007.

24. P.G. Alotto, C. Eranda, B. Brandstätter, G. Fürntratt, C. Magele, G. Molinari, M. Nervi, K. Preis, M. Repetto, and K.R. Richter, "Stochastic algorithms in electromagnetic optimization," IEEE Transactions on Magnetics, Vol.34, No.5, pp.3674-3684, September 1998.

25. P. Alotto and C. Magele, "Optimization in electromagnetics," International Compumag Society Newsletter, Vol.6, No.3, pp.7-9, 1999.

26. G.R. Walsh, *Methods of Optimization*, pp.97-102, John Wiley, 1975.

27. P.P. Silvester, "A general high-order finite-element waveguide analysis program," IEEE Transactions on Microwave Theory and Techniques, Vol.17, pp. 204-209, 1969.

28. S. Ahmed and P. Daly, "Waveguide solutions by the finite element method," Radio and Electronic Engineer, Vol.38, 1969.

29. J.P. Webb, "Application of the finite-element method to electromagnetic and electrical topics," Reports on Progress in Physics, Vol.58, pp. 1673-1712, 1995.

30. C. Johnson, *Numerical solution of partial differential equations by the finite element method*, Cambridge, U.K.: Cambridge University Press, 1987.

31. S.C. Brenner and L. Ridgway Scott, *The Mathematical Theory of Finite Element Methods*, New York: Springer-Verlag, 2nd Edition, 2002.

32. D.A. Lowther and P.P. Silvester, *Computer-Aided Design in Magnetics*, New York: Springer-Verlag, 1986.

33. J.C. Sabonnadiere and J.L. Coulomb, *Finite Element Methods in CAD: Electric and Magnetic Fields*, New York: Springer-Verlag, 1987.

34. E. Hardee, K.-H. Chang, J. Tu, K.K. Choi, I. Grindeanu, and X. Yu, "A CAD-based design parameterization for shape optimization for elastic solids," Advances in Engineering Software, Vol.30, pp.185-199, 1999.

35. ElecNet is a registered trademark of Infolytica Corp., *www.infolytica.com*, Montreal, QC, Canada.

36. Visual Basic and Microsoft Excel 2007 are registered trademarks of Microsoft Corp., USA.

37. D.M. Pozar, *Microwave Engineering*, New York: Wiley, 2nd Edition, 1998

38. R.E. Collin, *Foundations for Microwave Engineering*, New York: McGraw-Hill, 2nd Edition, 1992.

39. J.-M. Jin, *The Finite Element Method in Electromagnetics*, New York: John Wiley & Sons, 2nd Edition, 2002.

40. 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, 1999.

41. G.H. Golub and C.F.Van Loan, *Matrix Computations*, Chapter 10, Baltimore: John Hopkins University Press, 2nd Edition, 1990.

42. N.K. Nikolova, J.W. Bandler, and M.H. Bakr, "Adjoint Techniques for Sensitivity Analysis in High-Frequency Structure CAD," IEEE Transactions on Microwave Theory and Techniques, Vol.52, N0.1, pp. 403-419, 2004.

43. J.P. Webb, "Design sensitivities using high-order tetrahedral vector elements," IEEE Transactions on Magnetics, Vol.37, No.5, pp. 3600-3603, 2001.

44. 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, 1987.

45. 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, 1997.

46. M.A. Kolbehdari, M.Srinivasan, and M.S. Nakhla, Q.J. Zhang, R. Achar, "Simultaneous time and frequency domain solutions of EM problems using Finite Element and CFH techniques," IEEE Transactions on Microwave Theory and Techniques, vol.44, No. 9, pp.1526-1534, 1996.

47. X.-M. Zhang and J.-F. Lee, "Application of the AWE method with the 3D TVFEM to model spectral responses of passive microwave components," IEEE Transactions on Microwave Theory and Techniques, vol.46, No. 11, pp. 1735-1741, 1998.
48. J.E. Bracken, D.K. Sun, and Z. Cendes, "Characterization of electromagnetic devices via reduced-order models," Computer methods in applied mechanics and engineering, Vol. 169, nos. 3-4, pp. 311-330, 1999.
49. J.P. Webb, "Design sensitivity of frequency response in 3D finite element analysis of microwave devices," IEEE Transactions on Magnetics, vol.38, no.2, pp. 1109-1112, 2002.
50. W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, Numerical Recipes in C++, 2nd Edition, Cambridge University Press, 2002.
51. 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, 1999.
52. D. Nair and J.P. Webb, "Optimization of Microwave Devices using 3D Finite Elements and Design Sensitivity of the Frequency Response," IEEE Transactions on Magnetics, Vol.39, No.3, pp. 1325-1328, 2003.
53. V.A. Labay and J. Bornemann, "Generalized scattering matrix of waveguide corners distorted by discontinuities in the resonator region," IEEE MTT-S, vol.1, pp. 269-272, 1995.
54. 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, 1992
55. S. McFee and D. Giannacopoulos, "Optimal discretization based refinement criteria for finite element adaption", IEEE Transactions on Magnetics, Vol.32, No.3, pp. 1357-1360, 1996
56. D. Giannacopoulos and S. McFee, "Towards optimal *hp* adaption near singularities in finite element electromagnetics", IEEE Transactions on Magnetics, Vol. 30, No. 5, pp. 3523-3526, 1994.

57. P. Fernandes, P. Girdinio, M. Repetto, and G. Secondo, "Refinement strategies in adaptive meshing", IEEE Transaction on Magnetics, Vol. 28, No. 2, pp. 1739-1742, 1992.

58. O.C. Zienkiewisz and R.L Taylor, *The Finite Element Method*, Vol.1, 5th Edition, London: McGraw-Hill, 2000.

59. N.A. Golias and T.D. Tsiboukis, "Adaptive methods in computational magnetics," International Journal of Numerical Modeling: Electronic Networks, Devices and Fields, Vol. 9, pp. 71-80, 1996.

60. O.C. Zienkiewisz and R.L Taylor, *The Finite Element Method*, Vol.1, 5th Edition, Chapter 14, London: McGraw-Hill, 2000.

61. J. Van Bladel, *Singular Electromagnetic Fields and Sources*, New York: Oxford University Press, 1991.

62. D.N. Shenton and Z.J. Cendes, "Three-dimensional finite element mesh generation using Delaunay tessalation", IEEE Transactions on Magnetics, Vol. 21, No. 6, pp. 2535-2538, 1985.

63. F.G. Vier and O.A. Mohammed, "A 3-D finite element mesh generator for complex volumes", IEEE Transactions on Magnetics, Vol.30, No.5, pp.3539-3542, 1994.

64. J.P. Webb and B. Forghani, "Adaptive improvement of magnetic fields using hierarachal tetrahedral finite elements", IEEE Transactions on Magnetics, Vol.30, No.5, pp. 3511-3514, 1994.

65. I. Babuska and B.Q. Guo, "The *h*, *p*, and *h-p* version of the finite element method; basis theory and applications", Advances in Engineering Software, Vol.15, 1992

66. S. McFee and D. Giannacopoulos, "The implications of parallel processing on *h-p* adaptive finite element analysis for electromagnetics," IEEE Transactions on Magnetics, Vol. 35, No. 5, pp. 3284-3287, 1998.

67. W. Rachowicz, J.T. Oden, and L. Demkowicz, "Toward a universal *h-p* adaptive finite element strategy. Part 3: Design of *h-p* meshes," Computer Methods in Applied Mechanics and Engineering, Vol. 77, pp. 181-212, 1989.

68. J.T. Oden, T. Strouboulis, and P. Devloo, "Adaptive Finite Element method 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, No. 3, pp. 327-362, 1986.

69. D.W. Kelly, J.P.D.S.R. Gago, O.C.Zienkiewicz, and I. Babuska, "*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.

70. P.P. Silvester and R.L. Ferrari, *Finite Elements for Electrical Engineers*, 2nd Edition, Cambridge University Press, 1990.

71. O.C. Zienkiewicsz, 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.

72. D. Rodger, "Experience with hierarchal finite elements in 2D electromagnetics", IEEE Transactions on Magnetics, Vol.23, No.5, pp. 3560-3562, 1987.

73. J.P. Webb and B. Forghani, "Hierarchal Scalar and Vector Tetrahedra", IEEE Transactions on Magnetics, Vol.29, No.2, pp. 1495-1498, 1993.

74. J.P. Webb, "Edge elements and what they can do for you", IEEE Transactions on Magnetics, vol.29, No. 2, pp. 1460-1465, 1993.

75. G. Mur, "Edge Elements, their advantages and their disadvantages", IEEE Transactions on Magnetics, vol.30, No. 5, pp. 3552-3557, 1994.

76. L.S. Anderson and J.L. Volakis, "Adaptive multiresolution antenna modeling using hierarchical mixed-order tangential vector finite elements," IEEE Transactions on Antennas and Propagation, Vol.49, No.2, pp. 211-222, 2001.

77. J.-F. Lee, D.-K.Sun, and Z.J. Cendes, "Full-wave analysis of dielectric waveguides using tangential vector finite elements," IEEE Transactions on Microwave Theory and Techniques, Vol.39, No.8, pp/ 1262-1271, 1991.

78. M.M. Botha and D.B. Davidson, "The Implicit, Element Residual Method for a-Posteriori Error Estimation in FE-BI Analysis," IEEE Transactions and Propagation, Vol. 54, No.1, pp. 255-258, 2006.

79. M.M. Botha and J.M. Jin, "*A-posteriori* error indicators for 3D electromagnetics FE-BI Analysis," IEEE Antennas and Propagation Society International Symposium, Vol. 4, pp. 3497-3500, 2004.

80. A.B. Diaz-Morcillo, J.V. Balbastre, and L. Nuno, "New Recovery Error Indicator for Adaptive Finite-Element Analysis on Waveguide Structures," IEEE Transactions on Microwave Theory and Techniques, Vol. 51, No. 5, pp. 1467-1475, 2003.

81. W. Rachowicz and L. Demkowicz, "An *hp*-adaptive finite element method for electromagnetics-Part II: A 3D implementation," International Journal for Numerical Methods in Engineering, Vol.53, No.1, pp. 147-180, 2002.

82. 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, 1998.

83. D.-K. Sun, Z. Cendes, and J.-F. Lee, "Adaptive mesh refinement, *h*-version, for solving multiport microwave devices in three dimensions," IEEE Transactions on Magnetics, Vol.36, No. 4, pp. 1596-1599, 2000.

84. J.P. Webb, "P-adaptive methods for electromagnetic wave problems using hierarchal tetrahedral edge elements," Electromagnetics, Vol.22, No. 5, pp. 443-451, 2002.

85. D. Nair and J.P. Webb, "P-Adaptive Computation of the Scattering Parameters of 3-D Microwave Devices," IEEE Transactions on Magnetics, Vol.40, No.2, pp. 1428-1431, 2004.

86. 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," IEEE Transactions on Microwave Theory and Techniques, Vol. 45, No. 5, pp. 747-760, 1997.

87. M. Ainsworth and J.T. Oden, *A-posteriori error estimation in finite element analysis*, New York: John Wiley and Sons, 2000.

88. M. Salazar-Palma, T.K. Sarkar, L.-E. Garcia-Castillo, T. Roy and A. Djordjevic, *Iterative and Self-Adaptive finite elements in electromagnetic modeling*, Chapter 4, Boston: Artech House, 1998.

89. G. Drago, P. Molfino, M. Nervi, and M. Repetto, "A Local Field Error Problem Approach for Error Estimation in Finite Element Analysis," IEEE Transactions on Magnetics, Vol.28, No.2, pp. 1743-1746, 1992.

90. K.C. Chellamuthu and N. Ida, "*A-posteriori* Element by Element Local Error Estimation Technique and 2D & 3D Adaptive Finite Element Mesh Refinement," IEEE Transactions on Magnetics, Vol.30, No.5, pp. 3527-3530, 1994.
91. J. Penman and J. Fraser, "Dual and complementary energy methods in electromagnetism," IEEE Transactions on Magnetics, Vol. 19, No. 6, pp. 2311-2316, 1983.
92. J.A.H. Rikabi, C.F. Bryant, and E.M. Freeman, "An error-based approach to complementary formulations of static field solutions," International Journal for Numerical Methods in Engineering, Vol.26, No.9, pp. 1963-1987, 1988.
93. J.T. Oden and S. Prudhomme, "Goal-oriented error estimation and adaptivity for the Finite Element method," Computers and Mathematics with Applications, Vol. 41, pp. 753-756, 2001.
94. M.M. Gavrilovic and J.P. Webb, "An error estimator for design sensitivities of microwave device parameters," Electromagnetics, Vol. 22, No. 4, pp. 315-321, 2002.
95. R. Becker and R. Rannacher, "An optimal control approach to *a-posteriori* error estimation in finite element methods," Acta Numerica, Vol. 10, p. 1-102, 2001.
96. P. Ingleström and A. Bondeson, "Goal-oriented error estimation and *h*-adaptivity for Maxwell's equations," Computer methods in applied mechanics and engineering, Vol. 192, pp. 2597-2616, 2003.
97. J.P. Webb, "Using Adjoint Solutions to Estimate Errors in Global Quantities," IEEE Transactions on Magnetics, Vol. 41, No.5, pp. 1728-1731, 2005.
98. D. Nair and J.P. Webb, "Estimating Errors in Design Sensitivities," IEEE Transactions on Magnetics, Vol.42, No.4, pp. 559-562, 2006.
99. F. Alessandri, M. Dionigi, and R. Sarrentino, "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, 1995.
100. Y. Chung, C. Cheon, I. Park, and S. Hahn, "Optimal shape design of passive devices using FDTD and design sensitivity analysis," IEEE Transactions on Microwave Theory and Techniques, Vol.48, No.12, pp. 2289-2296, 2000.

101. Y. Chung, C. Cheon, I. Park, and S. Hahn, "Optimal design method for microwave device using time domain method and design sensitivity analysis-Part II: FDTD case," IEEE Transactions on Magnetics, Vol.37, No.5, pp. 3255-3259, 2001.

102. P. Arcioni, M. Bozzi, M. Bressan, G. Conciauro, and L. Perregrini, "Fast optimization, tolerance analysis and yield estimation of H-/E-plane waveguide components with irregular shapes," IEEE Transactions on Microwave Theory and Techniques, Vol.52, No.1, pp. 319-328, 2004.

103. E.A. Soliman, M.H. Bakr, and N.K. Nikolova, "Accelerated gradient-based optimization of planar circuits," IEEE Transactions on Antennas and Propagation, Vol.53, No.2, pp. 880-883, 2005.

104. M.A. El Sabbagh, M.H. Bakr, and J.W. Bandler, "Adjoint higher order sensitivities for fast fullwave optimization of microwave filters," IEEE Transactions on Microwave Theory and Techniques, Vol.54, No.8, pp. 3339-3351, 2006.

105. Y. Yang, T. Halleröd, D. Ericsson, A. Hellervik, A. Bondeson, and T. Rylander, "Gradient optimization of microwave devices using continuum design sensitivities from the adjoint problem," IEEE Transactions on Magnetics, Vol.41, No.5, pp. 1780-1783, 2005.

106. A. Bondeson, Y.Yang, and P. Weinerfelt, "Shape optimization for radar cross sections by a gradient method," Internation Journal for Numerical Methods in Engineering, Vol.61, pp. 687-715, 2004.

107. E.W. Weisstein "Triangle Inequality." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/TriangleInequality.html

108. D. Nair and J.P. Webb, "Accuracy-controlled combined adaption-optimization scheme for Finite Element analysis of 3D microwave devices," IEEE Transactions on Magnetics, Vol.43, No.4, pp. 1561-1564, 2007.

109. D. Nair and J.P. Webb, "An adaption-optimization scheme for improving the performance of microwave devices over a frequency range," 9th International workshop on Finite Elements for Microwave engineering, Bonn, May 8-9 2008.

110. C.G. Montgomery, R.H. Dicke and E.M. Purcell, *Principles of microwave circuits*, Chapter 9, New York: Dover publications, Inc., Chapter 1965.

111. J.W. Bandler, W. Kellerman, and K. Madsen, "A superlinearly convergent minimax algorithm for microwave circuit design," IEEE Transactions on Microwave Theory and Techniques, Vol.33, No.12, pp.1519-1520, 1985.

112. E. Polak, D.Q. Mayne, and J.E. Higgins, "Superlinearly convergent algorithm for min-max problems," Journal of Optimization Theory and Applications, Vol.69, No.3, pp.407-439, 1991.

113. R. Mittra, O. Ramahi, A. Khebir, R. Gordon, and A. Kouki, "A review of Absorbing Boundary Conditions for two and three-dimensional electromagnetic scattering problems," IEEE Transactions on Magnetics, Vol.25, No.4, pp. 3034-3039, 1989.

114. S. Groiss, I. Bardi, O. Biro, K. Preis, K.R. Richter, "Finite Element analysis of multiport fitlers using Perfectly Matched Layers," IEEE Transactions on Magnetics, Vol.33, No.2, pp.1480-1483, March 1997.

115. D.I. Karatzidis, T.V. Yioultsis, and T.D. Tsiboukis, "Gradient-based adjoint-variable optimization of broadband microstrip antennas with mixed-order prism macroelements," International Journal of Electronics and Communications, Vol.62, pp. 401-412, 2008.

116. Visual C++ 6.0 is a registered trademark of Microsoft Corp., USA.

117. Windows XP is a registered trademark of Microsoft Corp., USA.

Appendix A

The code implementation of the combined adaption-optimization software system was done using Visual C++ 6.0 [116] and the Visual Basic [36] scripting language, run within Microsoft Excel [36]. All codes were implemented on a Windows XP [117] platform installed on a machine with an AMD Athalon 1.67 GHz processor, with 1 GB of RAM and 40 GB hard disk space.

The combined adaption-optimization scheme has two main components: the steepest descent algorithm that uses the cubic line search interpolation technique, and the p-adaptive FE solver for computing the cost function and gradient.

The optimization algorithm is implemented in the Visual Basic scripting language. All stages of the line search algorithm, such as, the pre-optimization analysis step, the bounding analysis and cubic analysis step are all implemented within the script. The VB script runs the entire optimization, calling on the FE solver executable program only for the cost function evaluation process. The FE solver is implemented in Visual C++. The two components that comprise the p-adaptive FE solver, the error indicator and the error estimator, have both been implemented using Visual C++. Data files are exchanged between the two blocks in order to transfer relevant information pertaining to the analysis. The files are in the ASCII format that can be read by both the VB script (the optimizer) as well as C++ executables (the FE solver). A quick look at the main data files that are used for information exchange between the two blocks essential for the running of the combined scheme given by Figure 5.2 is the focus of this Appendix.

It must be noted that the relative accuracy term β from (5.1) is hard-coded into the solver, while the value of Γ, which is dependent on the frequency sweep values at each cubic point are updated on a regular basis throughout the analysis. Recall that the accuracy link formulation is dependent on the values β and Γ.

A.1 Data files output from the optimizer (input to the FE solver)

Information regarding the geometry generated from the latest design parameter vector needs to be passed on the FE solver for further processing, and eventually, to compute the cost function and the gradient.

A.1.1 bcs.dat

The boundary conditions (assigning PEC, PMC to the geometry) and port co-ordinate information (designating the excitation port, port orientation in the Cartesian co-ordinate system) required by the FE solver is generated based on information input by the user into an Excel worksheet. Each side of the geometry is identified and the proper boundary condition is assigned. The VB script generates the data file *bcs.dat* with this information which is passed on the FE solver.

A.1.2 tets.dat

Within the VB script, the user-specified design parameters for the given problem are used to create the desired geometry using the software package ElecNet. Once the model is created, ElecNet generates the appropriate mesh to approximate the geometry. The mesh data, which contains all the necessary mesh node information (such as the vertex co-ordinate information and model face-label information, explained in Section 3.1) are output to the file *tets.dat*. The information in this file is vital for the solver to assemble the universal matrix need for the FE solution. The information in this file is also crucial in identifying the nodes with their corresponding surfaces on the model when calculating the vertex sensitivities.

A.1.3 parameters.dat

This file passes on the information regarding the user-specified design parameters $\{g\}$ of the given problem. The information includes the number of design parameter values, as well as their values. Also included in this file are the values of the reference parameter vector {gStart}, as explained in Section 3.1. The mesh generation block of Figure 3.3 uses this data to determine the status of the perturbation process by evaluating the condition given by (3.1). The difference between {g} and {gStart} is used in conjunction with the vertex sensitivity values to determine the amount by which each node should be moved for mesh perturbation.

A.1.4 input.dat

This file contains the geometrical information needed for the vertex sensitivity calculations. The vertex sensitivity code is problem specific, whereby the main geometrical data of the problem (such as model dimensions, design parameter values) need to be input manually, rather than having the code automatically derive the values from the model. This information is needed when trying to match each mesh vertex with its corresponding surface, and then to calculate the sensitivities.

A.2 Data files output from the FE solver (input to the optimizer)

Information passed on to the optimizer from the FE solver includes the cost function and its gradient, as well as the status of the solution process itself, such as saturation conditions.

A.2.1 CdCdgVals.dat

This is the file in which the solver stores the cost function and gradient values. The values are output for calls to the FE solver from all stages of the optimization. The optimizer processes these values based on the stage that it currently is in, and takes the necessary action to advance the line search, as explained in Chapter 2.

A.2.2 Frequency.dat

The scattering parameter values calculated over the frequency range using the Pade approximation technique are output to this file. These values are mainly used for demonstrating the frequency response obtained for the latest design configuration.

A.2.3 log.dat

This file contains information regarding the solution process, such as the number of steps taken by the conjugate gradient solver to solve the global matrix and the number of non-zero elements in the system matrix. These values are used to estimate the computational cost involved in the solution process for a particular design point.

A.2.4 TolUpdateFlag.dat

The status of the FE solver is let known to the optimizer through a flag set by the solver after the analysis in this file. If the solver has hit saturation, the flag is set so that the optimizer is forced to terminate the entire optimization process.

A.2.5 gradCInfNorm.dat

The cost function gradient value from the frequency sweep required to calculate the accuracy term Γ is stored in this file. The optimizer reads in the value of $\Gamma^{(k)}$ for every cubic step analysis , but does not use it to update the accuracy link conditions (through Γ_{ref}) unless at the beginning of a line search. This was explained in detail in Chapter 5.