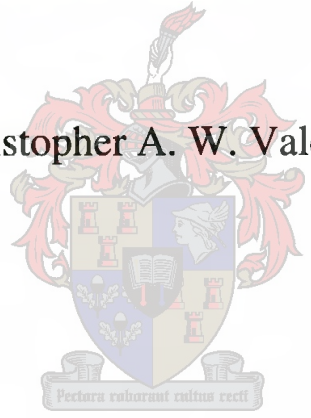
Growth-based Computer Aided Design Strategies for Multimode Waveguide Design with the Aid of Functional Blocks

Christopher A. W. Vale



Dissertation presented for the degree of Doctor of Philosophy in Engineering at the University of Stellenbosch.

Promoter: Prof P Meyer

Date: December 2001

Declaration: *I, the undersigned, hereby declare that the work contained in this dissertation is my own original work, and has not previously in its entirety, or in part, been submitted at any university for a degree.*

Signed:

Christopher Vale

Date: 28 - 11 - 2001

Abstract

Key words: *CAD, Overmoded Waveguide, Multimode Waveguide Devices, Mode-Matching, Measurement, Multimode Horns*

A new technique for the design of multimode devices in overmoded waveguide is presented. The technique applies the principle of growth-based design and uses a conceptual functional block representation of the design structure to provide necessary flexibility to the design algorithms. Two growth based design strategies are proposed and evaluated. The first uses a generalized synthesis-oriented scanning technique, and the second uses an evolutionary strategy. The techniques provide reliable solutions to a variety of multimode design problems.

In order to facilitate sufficiently fast numerical analysis, novel enhancements of the mode matching technique are developed and the use of surrogate models is investigated. In addition, to allow physical evaluation of the finished devices, original techniques of measuring multimode devices are formulated and utilised.

Two practical problems are used to evaluate the performance of the design procedures. The first is the design of overmoded waveguide chokes for microwave heating facilities, and the second is the design of multimode horns for antenna and spatial power combining applications. Various examples of each type of problem are presented with measurements of manufactured solutions.

Opsomming

Sleutelwoorde: *CAD, Multimodus Golfleier, Multimodus Golfleierstelsels, Modale-Pas Tegnieke, Meting, Multimodus Horings*

'n Nuwe tegniek vir die ontwerp van multimodusstelsels binne multimodus golfleier word voorgestel. Die tegniek maak gebruik van die beginsel van groei-geöriënteerde ontwerp en ontgin 'n konsepsuele funksionele module-voorstelling van die ontwerpstruktuur om die nodige buigsaamheid aan die ontwerpsalgoritmes te verleen. Twee groei-geöriënteerde ontwerpstrategieë word aangebied en geëvalueer. Die eerste is gebasseer op 'n veralgemeende sintese-geöriënteerde skandeertegniek, en die tweede maak gebruik van 'n evolusie-strategie. Die tegniek verskaf betroubare oplossings vir 'n verskeidenheid van multimodusontwerpsprobleme.

Ten einde 'n numeriese analise-tegniek daar te stel wat vinnig genoeg is, word oorspronklike verbeterings van die modal-pas metode ontwikkel en surrogaatmodelle is ook ondersoek. Verder, vir fisiese evaluasie, word oorspronklike meettegnieke vir multimodusstelsels geformuleer en gebruik.

Twee praktiese probleme word gebruik om die ontwerpprosedures te evalueer. Die eerste is die ontwerp van multimodus golfleierdrywingsdempers vir mikrogolfverhitting, en die tweede is die ontwerp van multimodus horings vir antenna- en ruimtelike drywingskombineerder-toepasings. Verskeie voorbeelde van elke tipe probleem word gegee met metings van geïmplementeerde oplossings.

Acknowledgements

This thesis was made possible by a number of contributions from various people. My sincerest gratitude and thanks to my promoters, Prof. Petrie Meyer and Prof Keith Palmer for their many insights and suggestions and wholehearted support of this work.

I would also like to acknowledge the National Research Foundation for supporting this research financially and providing me with the opportunity to attend various symposia.

Many thanks also to Wessel Croukamp and the SED for the manufacture of a variety of tricky designs and to Dr. J.B. de Swardt for the excellent photographs. I would like to acknowledge the Microwave Dielectric Heating Group, specifically Prof. H.C. Reader and Mr. M Rimbi for allowing me to participate in and share the results of the design of a microwave heating choke.

To the inhabitants of room E251, Werner Steyn for many valuable conversations, Thomas Stuart for all the computing power, Mark Rutschlin and Martinette Muller, thanks for making this effort an enjoyable one.

Most especially to my loving father and mother and my dear friend Ellen Louw. Thank you for your love and support, and for never doubting, even when I did.

Finally, I thank God for giving me the strength and granting me the insights needed to complete this work.

Contents

Chapter 1

Introduction

1.1 Design Problems in Overmoded Waveguide

When waveguide is excited at sufficiently high frequencies, it becomes capable of supporting multiple propagating modes. This type of behaviour is generally avoided by designers as it introduces a number of complications that make it very difficult to apply standard design techniques. Despite these inherent complications, multimode waveguide devices currently enjoy many diverse applications, and with the increasing demand on designers to produce ever cheaper, more efficient and more complex designs, these applications can be expected to increase in the future.

The first applications of multiple propagating modes in a single device were in the feeds of reflector antennas, where a number of propagating modes were used in combination to build up desired antenna field patterns. One of the first uses of modes in this way was proposed by Potter as early as 1963. This application made use of two orthogonal modes at the aperture of a conical horn to improve the horn's performance characteristics [Potter 1963]. Since then, multimode horns have grown in popularity and a number of design approaches have been proposed. These devices will be discussed in more detail in chapter 6. The emerging field of spatial power combining also makes use of multimode horns, only not to radiate into half space, but rather to distribute energy evenly over an array of active devices [Harvey 2000]. Similar potential uses exist for field-shaping with multiple modes, including low loss transmission of microwave energy and selective illumination of other types of devices in waveguide.

Microwave power generation devices typically incorporate overmoded waveguide sections to reduce ohmic losses and to avoid breakdowns. In plasma research, mode converters are often used to convert high power modes, used for transmission in such systems, to the fundamental modes used for research purposes [Buckley 1990, Lunéville 1998, ul Haq 1998]. A concern

with such devices is maintaining compactness whilst ensuring a good conversion efficiency and acceptable power handling. Low loss twists in overmoded rectangular waveguide have been developed for power transmission in order to overcome unfavourable bends separating receiver and transmitter devices [Doane 1988]. More recently, Rosenberg et al. proposed a CAD method based on mode matching for the design of high performance transitions from rectangular to overmoded elliptical waveguide [Rosenberg 2000]. The technique makes use of a modified Newton method to optimise the dimensions of a rough starting transition comprising a number of transition elements in cascade.

In some situations, mechanical requirements have enforced the use of overmoded waveguide, leading to some complex design issues. At very high frequencies, mode filters have been developed to suppress higher order modes in sub-millimetre wave applications where waveguide dimensions cannot be fabricated small enough to support single mode propagation [Butterweck 1968]. In microwave dielectric heating, chokes are designed to suppress the propagation of modes in a variety of typically overmoded waveguide geometries. Various design approaches have been proposed, including the use of corrugated filter geometries [Soto 2000]. Typically, however, the proposed strategies must enforce constraints on the choke aperture geometries, limiting their general application. The design of these devices will be discussed in more detail in chapter 5.

1.2 Problems with Classical Design Techniques

The mainstay of passive microwave design lies with the principle of network synthesis, developed by Darlington and others from as early as 1939. This principle is typically exploited to obtain lumped and/or distributed element circuits that realize a prescribed frequency response that is described by a rational polynomial. Work done by many in the field, notably Levy, Richards and Matthaei, allow such lumped element circuits to be transformed (with, for example, impedance inverters) to forms that can be more readily applied to microwave structures. The final step of design attempts to recreate the circuit in a microwave transmission medium, such as waveguide or microstrip, by mirroring the electrical nature of each individual circuit element with microwave structure, like an iris, cavity, or length of waveguide. For many years, the *Waveguide Handbook* [Marcuwitz 1951] provided

the pre-eminent mapping between such physical structures in waveguide and the idealized circuit components required by synthesis. The modern approach is to rely on more accurate and flexible numerical modelling techniques to characterise microwave structures. These techniques have allowed more varied and unusual structures, such as dielectric pills, to be used as elements in microwave circuits.

The underlying philosophy, however, is the solution of a simpler, circuit level design problem, which can often be done using analytical techniques based on synthesis, and the subsequent hunt for microwave substructures that can mimic the circuit components. The application of CAD in conjunction with this strategy is widely covered in the literature and will be discussed in more detail in the next section. While this general approach is a powerful and eminently applicable strategy, there are a number of complications that can hinder its effectiveness or even make its use impossible.

a) Modelling limitations

Circumstances can arise where the type of microwave substructure used cannot be modelled by a simple lumped element equivalent, or where the structure's model is too involved to be of any use by the circuit-level design technique. This is especially prevalent with multimode discontinuities which have to be modelled as multiport scatterers, typically with complex frequency responses and a degree of cross-coupling between the propagating modes.

Bandwidth can also be a problem. Typically, sufficiently simple lumped and/or distributed element approximate circuits are accurate for only a limited bandwidth around the design frequency. Standard design applied to high bandwidth problems using such band limited models can cause serious inaccuracies in design, resulting in a device that performs acceptably for only a very limited part of its working bandwidth.

b) One device – Multiple functions

Synthesis can be very difficult to apply directly in situations where a single physical structure has to perform more than one electrical function. This is because synthesis is formulated in terms of two-port networks, and is geared towards obtaining a physical realization from a single response specification. For a device with multiple electrical specifications, each response will yield a different physical device with synthesis. It is then up to the designer to

find an intelligent way of combining the devices in one physical structure. In situations where each electrical port corresponds to a physical port, such as in a simple diplexer, this can usually be accomplished (though often with some difficulty) by cascade or parallel arrangements of the individual devices.

In multimode devices, where the ports are not physically isolated, the modes are forced to see the same physical structure. However, since each mode experiences the structure differently on a circuit level, it is difficult to control a response at one of the ports without disturbing the response at another port. This problem is additionally compounded when cross-coupling occurs between the modes, effectively mixing different responses with each other. This type of behaviour is easily observed in single mode filters at sufficiently high frequencies, where the frequency response deviates from that predicted by circuit theory, unless discontinuities are carefully chosen so as not to induce cross-coupling.

c) Impossibly high performance requirements or unusual specifications

Various topologies exist for realizing passive components in waveguide and other media. Their diversity is, in part, due to the ever present need for high performance solutions. Since the number of topologies is finite, it is, however, conceivable that for a particularly stringent design, no known topology may suffice, leaving the designer with no option but to attempt to develop a new topology that can be used in conjunction with standard synthesis. Equally problematic are desired responses or specifications that simply cannot be treated as variations on the well-known filter problems. These include responses that are not completely specified or are specified in a way that cannot be directly related to a circuit problem.

d) Outside influences enforce limitations on the design

Synthesis-based design relies on the designer choosing a circuit topology that is of such a nature that, when realized as a design, it satisfies not only electrical specifications, but also peripheral requirements, such as maximum weight or length. As before, with a finite number of available topologies, it is possible that the outside influences could exclude all available topologies, leaving the designer with no way of applying synthesis to solve the design problem.

When problems exhibit one or more of the types of complications mentioned above that preclude the use of standard network synthesis-based techniques, the designer is left with no framework upon which to construct a solution. As a result, when these types of problems are encountered, a great deal of work and intellect is typically spent finding ways of either modifying the problem in some way, so that it can be solved with circuit synthesis, or of solving the problem without the aid of circuit synthesis at all.

Methods for solving such problems are many and varied. A common approach is simply to allow only for very special cases of the general design solution which can be tackled by conventional techniques. Examples of this are the design of hybrid couplers, where symmetry is enforced to allow the use of standard two-port techniques [Reed 1956]. A second example is the design of waveguide chokes for microwave heating applications, where the choke aperture geometries must be restricted to allow for mono-mode synthesis techniques [Metaxas 1983]. When such specializations are not possible, designers often resort to time-consuming trial and error techniques, which must typically be combined with a sizeable measure of knowledge and intuition regarding electromagnetic behaviour if they are to produce workable results. Other approaches include working from design graphs obtained by other researchers through physical experimentation, approximating wave behaviour to obtain a rough starting point for optimisation, and breaking up the design into sub-problems which can be tackled individually before subsequent re-incorporation and lot of optimisation.

Recent years have seen a broadening of computer aided design's role from that of being a useful tool to support what was essentially synthesis-based design, to a fully-fledged alternative design methodology, useful when designers are faced with problems of the types mentioned above. This approach uses computers to develop design solutions independently. These strategies typically make use of iterative optimisation methods to search design spaces for viable solutions to elusive problems. The nature of this computer aided design based approach will be discussed in the next section.

1.3 Applications of Computer Aided Design (CAD)

CAD applications can be divided into two classes. The first typically uses CAD to aid in the refinement of finished designs and for the identification and tuning of substructures corresponding to lumped element equivalent circuit elements. In this capacity, it is a secondary design tool, and does not drive the design itself. The second class of application, however, uses typically iterative CAD procedures to develop the design from little or no starting point, usually in the absence of a suitable deterministic design procedure.

The first class of application occurs more frequently in the literature. It dates back as far as 1967 to a paper by Temes and Calahan discussing the use of optimisation in circuit design [Temes 1967]. A detailed history of its development for microwave and millimetre wave applications, due to Bandler et al., indicates that research up to the late 1980s was primarily concerned with increasing the efficiency of optimisation procedures through the improvement of existing optimisation algorithms and the development of such breakthroughs as the adjoint network method, which allow much faster determination of objective function gradients [Bandler 1988]. Optimisation was applied in the CAD environment both to account for modelling approximations in finished designs, as well as for modelling circuit structures and designing for maximal resistance to manufacturing tolerances. The 1990s saw CAD widely being used to acquire equivalent microwave structures for lumped element components which were originally determined through the use of synthesis [Arndt 1997]. A novel design technique, proposed by Guglielmi, differs from this general theme by progressively matching the partially realized structure with the equivalent partial network representation [Guglielmi 1994]. This strategy offers some advantages in minimizing the amount of total optimisation that must be done to realize a suitable solution. In a number of applications, numerical analysis was tied to optimisation for the design of a variety of components from rough initial starting points [Alos 1997, Alessandri 1997, Alessandri 1995]. The last of these contributions is notable owing to its direct use of the adjoint network method in the optimisation routine. Direct optimisation was used with some success in the design of coupled cavity filters as described by Arndt et al. [Arndt 1997]. In this application, synthesis was avoided by directly optimising an appropriate equivalent circuit to achieve a desired (4 pole) response before the design was transformed to a microwave structure. Although this design did not make use of synthesis, the specification of the equivalent circuit requires a degree of prior knowledge

regarding the filter's coupling structure. These types of strategies are summarized in a good description of the state of the art by Gupta, where it is also indicated that a growing importance in CAD is ensuring accurate modelling of millimetre wave components, due to the prevalence of parasitic couplings at higher frequencies [Gupta 1998].

When CAD is applied in the ways mentioned above, optimisation is generally combined with numerical analysis either to refine designs that under-perform owing to inaccuracies in modelling or to aid in the identification of microwave structures for equivalent network realization. Other applications of conventional CAD are for design centering, to make provision for manufacturing tolerances, or even, in a limited way, for adjusting physical variables to an external requirement. The types of optimisers used in conventional CAD reflect its nature; they are typically highly efficient gradient based optimisers, suited to effective location of the closest local minima in the fewest possible iterations. Their performance is consequently related to the proximity of the starting structure to the global minima.

In the absence of a rough initial starting point for a design, or when circuit level specifications cannot be used to drive synthesis, CAD driven design offers a alternative design methodology that shows a greater flexibility than synthesis-based design, but an increased demand in terms of computational resources. In many cases it may be the only option available. This usually requires a more sophisticated approach than direct optimisation alone, as the outcomes of direct optimisation typically depend on the starting position and on the number of local minima in the error function. Any design that cannot be treated analytically will also be sufficiently complex to derail most attempts at direct optimisation. Furthermore, the large number of design variables typically associated with such problems make exhaustive search procedures impossible.

Successful design techniques generally exploit knowledge of the problem, or use a problem-specific shape optimizer. A principle that has been used with some success is the gradual incorporation of complexity into the solution as the design progresses. This tends to steer the solution to a globally satisfactory result. This type of strategy has been successfully employed in the computer aided design of finline tapers [Vale 1999] and mode converters [ul Haq 1995]. Both these examples resulted in designs that could not be matched by standard

techniques. Evolutionary strategies have also been used in a similar way with some success in the design of analog non-linear circuits [Koza 1997]. Their cost, however, remains excessive computational requirements that can limit the scope of application of the procedure.

At present, the number of techniques for this type of CAD are still very limited, and have the following problems:

- CAD driven design procedures are usually less reliable than deterministic procedures in terms of the repeatability of design outcomes. Typically, a number of different solutions of varying quality can result from repeated application of such techniques to any single problem.

- The iterative nature of such techniques, combined with the necessity to make use of numerical field solvers to account for complex structures, can often lead to excessively long design times for these types of techniques. Reducing the numerical analysis time is, consequently, a critical factor in making CAD-based techniques viable.

- Though CAD is appreciated for its flexibility in terms of being able to deal with a variety of problems and specifications, most successful CAD procedures are formulated for specific problem types, using a large degree of problem-specific information. This implies that though CAD is highly applicable, a new procedure or piece of software must be written for each type of problem, limiting its attractiveness for general use.

1.4 A New Approach to Multimode CAD

This dissertation proposes a new approach to the computer aided design of multimode devices in rectangular waveguide using the principle of growth-based design strategies with the aid of functional blocks .

This technique first determines a class, or number of classes of waveguide discontinuities that, when used together, are most probably able to satisfy a specific set of design requirements. As a group, these discontinuities are termed functional blocks. A structure, formed by cascaded

functional blocks, is then grown by gradually increasing the number of functional blocks in an intelligent way, i.e. by using a growth-based strategy.

Two example problem classes are addressed, both resulting in original solutions, with improved performance when compared to existing solutions. One is the design of overmoded bandstop filters for microwave heating applications, and the other is the design of multimode horns for antenna and spatial power combining applications.

The new growth-based, functional block CAD procedures and the two applications form the main contributions of this work. In addition, a number of smaller contributions are also presented. These are:

Improvements in analysis techniques

- A mode matching-based equivalent circuit for general waveguide discontinuities.
- An algorithm for intelligent mode selection in the mode matching procedure, resulting in considerable reductions in the number of modes that have to be considered for general cascaded systems and consequent increases in analysis speed.
- The incorporation of surrogate models in the design procedures, providing further reductions in design time if so desired.
- Modal transforms that allow conversion between scattering parameters expressed in terms of different modal sets in rectangular waveguide.
- A solution to the problem of ill-conditioning in T-parameter cascading of general discontinuities in waveguide.

Improvements in measurement techniques

- A technique that allows the determination of the generalized scattering matrix of a multimode device with minimal effort through measurement with a standard two port network analyzer.

1.5 About the Thesis

The next chapter will introduce the proposed growth-based CAD strategy devised to address the problem of multimode design. It will provide a general motivation for this choice of approach as well as detailed descriptions of the design algorithms.

Chapter 3 details the development of analysis techniques to apply with the CAD strategies developed in chapter 2. A number of original contributions that provide gains in analysis speed and problem insight will be detailed.

Chapter 4 details the development of novel measurement techniques used to evaluate the solutions to overmoded design problems. The techniques developed represent an important aspect of multimode design, not only allowing the validation of the numerical analysis techniques developed in the previous chapter, but also allowing the designs to prove themselves under practical conditions.

Chapter 5 introduces the first application of the CAD strategies to a practical design problem, notably in the design of overmoded chokes for microwave heating applications. The design approach proposed represents an original, published contribution to the field and is evaluated in a practical application.

Chapter 6 tackles the problem of multimode horn design. As in the previous chapter, the strategies developed in chapter 2 are successfully applied to develop novel design solutions for horns in antenna and spatial power combining applications. The design solutions are evaluated by measurement and comparison with other designs presented in the literature.

Chapter 7 concludes the thesis.

Chapter 2

Growth-based Computer Aided Design Strategies with Functional Blocks

2.1 Introduction

The wide range of design problems encountered in multimode applications encourages the development of design tools that are sufficiently flexible to cope with varying design demands. Not only are the electrical specifications likely to be of a changeable nature according to the systems in which the devices are to be used, but complex physical constraints might have to be adhered to that greatly affect the outcomes of the designs. Other considerations are the reliability and efficiency of the design procedures. The methods developed should provide assurance of a usable design solution without the expenditure of excessive computational effort.

The development of CAD techniques which conform to these requirements demands not only an extensive analysis of the nature of the problem and of any existing techniques that might be applicable, but also calls for the development of a conceptual method of dealing with the physical structures to be designed. The first part of this chapter motivates the choice of growth-based CAD strategies for the design of multimode waveguide components and suggests a conceptual model, based on functional blocks, that can be used in conjunction with the CAD strategies to manage the solution structures. This is followed by a discussion of the types of specifications and constraints likely to be encountered and how they will be accounted for in the various design strategies. The following section will detail the functional block-based method of structure characterization and explain the mechanisms it employs to ensure that physical structures adhere to their many constraints. The last part of this chapter is devoted to fully explaining and characterizing the Growth-based CAD strategies themselves.

2.2 The Problem of Multimode Design

The structure considered in this dissertation is waveguide with a single physical input port and a single physical output port, each potentially supporting a number of possible modes. The solution structure would typically consist of a series of discontinuities of arbitrary nature, separated by lengths of uniform guide. The problem is one of finding the appropriate arrangement and nature of the discontinuities, as well as the lengths of line separating them.

Design in this environment suffers from a number of complications. Some of these complications are due to the choice of structure itself, and some are due to the way the desired electrical responses are specified. Understanding of these complications is essential to the formulation of a viable design procedure, as it highlights the essential difficulties from which existing design techniques suffer.

Complications due to structure arise primarily because of the simultaneous existence of a number of electrical systems in only one physical structure. Consider the simple case of a length of waveguide line. In multimode waveguide, a single length of physical waveguide will appear to have a different electrical length for each mode it supports. Because of this, the simple design of an overmoded quarter wave transformer can become completely unworkable. To make matters more complicated, the relationship between the different electrical lengths is also variable, depending on both the dimensions of the waveguide and the frequencies at which the modes are propagating. If it were essential for a particular phase relationship to be enforced between two modes over a length of multimode waveguide, it might be possible to find an appropriate discontinuity that could be incorporated in the waveguide line so as to provide phase lag to one mode and not another. The effect that this discontinuity would have on other propagating modes, however, would have to be taken into account before this could be safely implemented. There is the possibility that such a discontinuity might cross-couple energy between the modes, causing unwanted electrical effects further on in the device. Clearly, the inclusion of any element, whether it be a simple discontinuity or a length of line, has to be carefully considered for all the modes in the system. The difficulty in modeling discontinuities stems largely from complications arising from the structure. The number of electrical roles that an overmoded discontinuity plays must be reflected by a sufficiently complex model. Design, however, typically requires simple

models, with electrical elements that can be independently controlled by the physical variables. In an overmoded environment, such independent control is difficult to achieve. This implies that design cannot make demands of individual electrical structures in the same way that it can in single mode design, since there is no guarantee that all the electrical structures will all be realizable by one discontinuity simultaneously.

A second complication of multimode design lies with the way the designs are specified. The theory of passive networks provides standard two-port design with the powerful knowledge of what is possible. This allows exact specification of desired responses and, consequently, the ability to pursue a planned design strategy. As yet, no mechanism exists that allows the exact polynomial-style specification of multimode devices. As a result, multimode device specification must be satisfied with secondary requirements that relate typically to the application being pursued. Such specifications may demand certain relationships between electrical parameters, but not specify them directly, or, otherwise, they might not be directly concerned with certain electrical properties such as phase. The difficulty in applying planned design strategies without full characterization of the desired structure arises because of a lack of information in the specification. Planned design strategies are effectively inverse problem solvers – given an outcome they can solve directly for the cause. Under the right circumstances, this is a powerful approach, but if the outcome is not completely specified, the solver is rendered ineffective.

This investigation examined a number of existing design techniques in pursuit of a general approach that might be applicable to multimode design. The various approaches can be broadly classified as either planned, synthesis-style approaches, or unplanned, optimisation-style approaches. The complications of applying classical, planned synthesis approaches to this type of problem have already been discussed in chapter 1 as a motivation for the development of alternative CAD-based techniques. Accordingly, this section will focus on alternative planned approaches found in the literature.

A technique proposed by Tascone et al. offers some promise for multimode design as it avoids the need to model discontinuities as equivalent circuit elements by pursuing synthesis directly in distributed media [Tascone 2000]. The technique synthesizes single mode filters as equal lengths of transmission line separated by discontinuities. The discontinuities are

specified by their desired scattering parameters, making it easier to implement the design directly in distributed media. Application of this technique to multimode design, however, suffers from the problems typically encountered in planned design. First, the technique requires full definition of the device T-parameters. In the single mode case, this is done by exploiting FIR digital filter responses, which are of an equivalent mathematical nature to the responses realized by the simplified T-parameter based models used in this technique. Furthermore, despite the fact that the technique avoids the difficult issue of equivalent lumped element modeling, it still demands a fully specified set of scattering parameters for each discontinuity, which cannot be guaranteed in the multimode environment. An additional complication is that when multimode lines separate the discontinuities, the variable electrical lengths seen by the modes greatly complicate the simple T-parameter based model that this approach relies upon, making it very difficult to perform the task of element extraction. A last complication is that the T-parameter based approach does not lend itself well to different port sizes. In multimode design, a smaller port which supports fewer propagating modes, is typically represented by fewer electrical ports than a large physical port.

The design of microwave filters by inverse scattering was first proposed by Roberts et al. in 1995. The approach makes use of a transform applied to the telegraphists equations that cause them to obey the one dimensional Schrödinger equation. Once in this form, inverse scattering can be applied to the fully specified reflection coefficient to obtain the desired impedance profile of the filter [Roberts 1995]. Although the method is extremely powerful, it is oriented specifically at planar circuit filter realizations and would be difficult to implement in an environment comprising waveguide discontinuities. The application of Inverse Scattering to the problem of filter design, despite being an interesting variation of planned design, still requires a complete specification of the desired response for its successful application.

De Pádua et al. present a similar technique for planar linear filter realization based on the inversion of the TLM method [de Pádua 2000]. This approach lends the technique a higher level of accuracy, but still relies upon the complete specification of the desired reflection coefficient of the device. In addition, it is questionable whether its application in the waveguide environment will be viable.

Those approaches that pursue unplanned design generally use iterative techniques to feel their way to a viable design. As outlined in chapter 1, these techniques typically rely upon a good starting point for the design. Such starting points do not come easily in multimode design owing to the general lack of knowledge about the media and its inherent complications. Furthermore, these approaches are incapable of developing new topologies, thereby making it impossible for them to create new designs that are potentially far superior to those already in existence.

2.3 A New Design Strategy

Synthesis-based design attains success by gradually adding elements to a design in such a way as to make it progressively more capable of fulfilling its function. Instead of creating a design solution in an instant, synthesis extracts the element that best simplifies what remains to be designed.

Working on the premise that multimode synthesis would be possible if the full multimode device specification were available and the multimode elements were indeed susceptible to modeling, it is possible to identify a synthesis oriented approach to multimode design that mimics the property of synthesis to build up solution structures. This approach attempts to add elements successively to a multimode structure with the intention of producing a new structure which is ever more similar to the design solution. Owing to the lack of a full specification, however, the only feedback that such a scheme could rely upon is how much closer the structure incorporating the newly added element is to the desired specification than to the original structure. To identify the most appropriate element for addition to the structure, all of the available elements would have to be tested. That element most likely to have been placed by a real synthesis approach would then be incorporated in the design. Without any other information being available, it seems reasonable to conclude that the element that causes the greatest improvement in the structure would be the most sensible element to include in the structure.

Following the standard synthesis approach, where the structure is built up by applying the elements at one end of the emerging design, would, however, not account for the necessary

inclusion of elements that may not initially appear to provide improvement in the design. Such elements might be left out and have no means of being incorporated at the correct location at a later stage of the design. To account for this, the strategy would have to allow the inclusion of elements at arbitrary points in the structure.

The emerging design strategy can be broadly classified as a growth-based approach, as it gradually allows the complexity of the solution structure to increase as the design progresses. There are numerous advantages of such a scheme over the cruder approach of starting with a high level of complexity and attempting to resolve an ordered design from it. Most notable is that the level of complexity incorporated in the final solution will be appropriate to the degree of difficulty of the problem. In addition, for multimode designs, where solutions might demand slight variations to account for small discrepancies between the modes, such small and secondary design aspects are automatically added a later stage of the design and do not complicate the overall structural development at early, formative stages. The net effect of starting with lower complexity is that the design will be steered towards a global design optimum, rather than be distracted by the resolution of small details of design. In keeping with the general approach of standard waveguide design, and to ensure that fabrication of potential solution structures is viable and relatively easy, it is proposed to apply the growth-based approach to the development of structures consisting of individual discontinuities separated by lengths of waveguide. This is chosen over other possible topologies, such as smoothly continuously varying profiles, or corrugated structures, as it not only allows a more simplified analysis of the structures using well known techniques, but allows a more discrete measure of growth. Since the discontinuities are likely to perturb the modal propagation, their inclusion will introduce complexity into the structure: the more discontinuities, the more inherent complexity in a structure.

Since such a wide range of possible discontinuities exists in waveguide, a technique must be developed that allows the seamless incorporation of various different types of discontinuities in a structure. The inclusion of a new discontinuity should, if at all possible, not overly perturb the rest of the structure and should not require special treatment from the design procedure. It is, therefore, proposed that design be implemented on a conceptual level where the discontinuities are considered as functional blocks and their physical details are masked from the design procedure. Design implemented on this level can be applied to a number of

different problem types without restructuring of the CAD algorithms, providing re-usability and ease of development.

Taking both aspects into account, a new design approach emerges, that of a growth-based algorithm, implemented with the aid of functional blocks.

2.4 Describing Structures with Functional Blocks

The approach discussed above calls for a procedure that allows the arbitrary addition of excitation elements in a solution structure whilst maintaining the general shape of the structure to the greatest possible extent. It also calls for the ability, if so desired, to modify, or even remove excitation elements or separating line sections already in the structure. The excitation elements, or *functional blocks*, each conceptually represent a type or class of discontinuity and are separated from adjacent functional blocks by a section of uniform waveguide. A functional block need not be limited to represent a simple discontinuity. It may, if so desired, also represent a more complex substructure, such as a resonator.

To make the design strategies as generally applicable as possible, they should not have to account for the various types of functional blocks on an individual basis - a functional block should merely be represented by a series of **design variables** at a specific location in the structure. The addition of a functional block should, in turn, be represented by the addition of a new series of design variables inserted at some point in the general structure. To achieve this level of flexibility, the design strategies are developed to pursue design on a simple conceptual level, shown in figure 2.1, where the structure is made up of a system of generic functional blocks and line sections, each with its own design variables.

Before analysis occurs, a mapping to the real physical structure is determined by the 'Rubberband Algorithm', so termed because its function is to stretch the physical structure to the demands of the design variables as closely as possible. This typically entails the assignment of various values to physical dimensions and the determination, sometimes through optimisation, of a viable method of connecting the functional blocks with waveguide line sections.

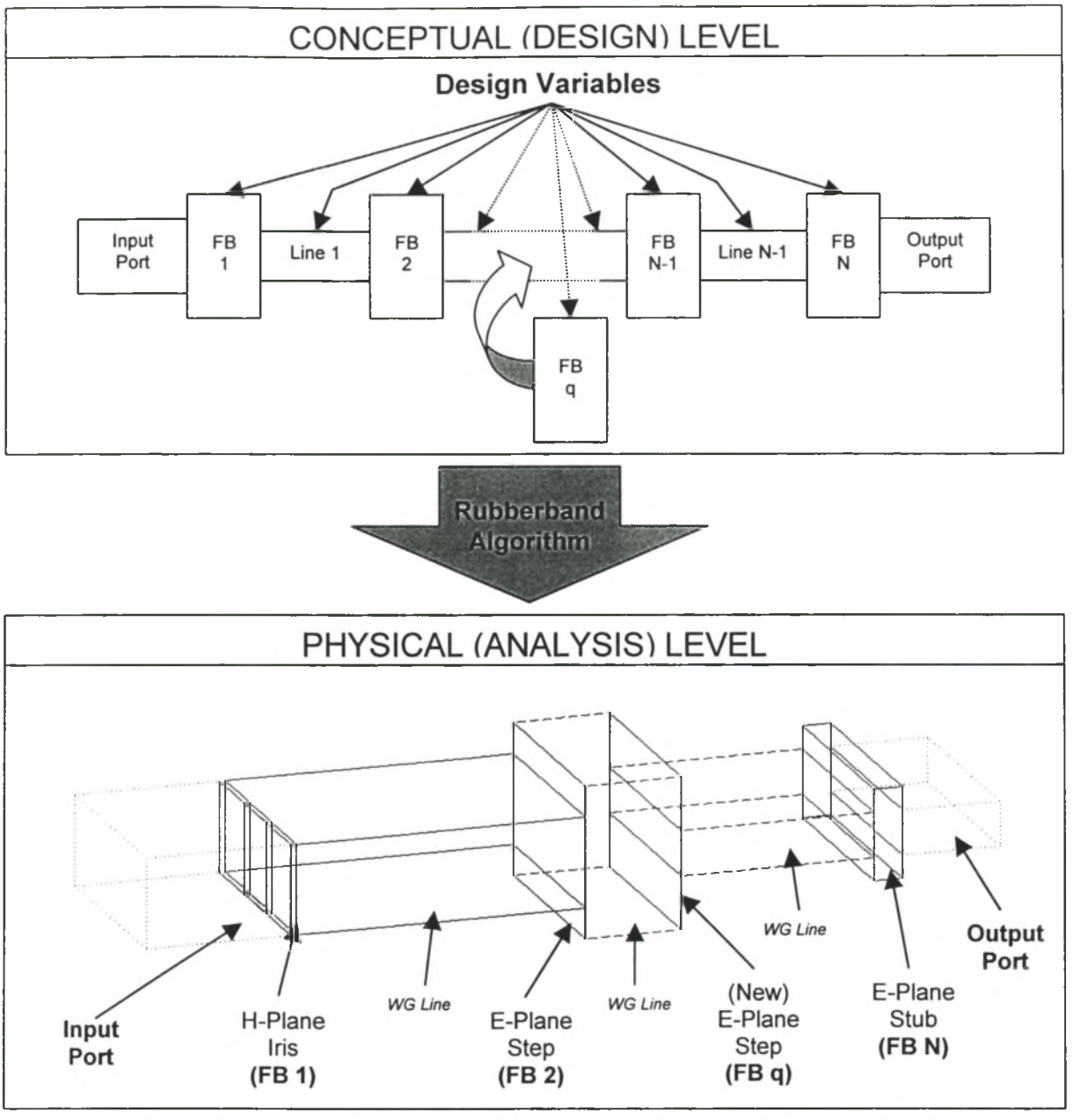


Figure 2.1 The general layout of a structure in terms of functional blocks

2.4.1 The Functional Blocks

A functional block consists of any discontinuity with an input and output port that can be coupled to a rectangular uniform waveguide of arbitrary size. It is defined by a number of **structural variables** that uniquely describe its dimensions. Figure 2.2 shows the structural variables of a general H-Plane Iris as implemented for this investigation.

Structural variables are either specified as *changeable* or *fixed*, according to the preferences of the designer. Any variable that is fixed will retain the value it is given by the designer throughout the entire design procedure for all instances of the functional block used in the structure. If, for example, a design incorporating the functional block in figure 2.2 demanded that all iris apertures be set at a width of 0.1mm owing to a desired choice of fabrication, the variable corresponding to this physical attribute could be set as fixed at this value for the duration of the design process.

Changeable structural variables can have different values for each instance of the functional block in the structure. As they are the parts of the structure subject to design, their values are influenced by the functional block's design variables. In the example above, a design variable might, for example, be defined as the aperture width, implying that the Rubberband algorithm has to perform a direct mapping from design variable to structural variable. Such a trivial mapping is often, but not always, the case, and depends on the way the functional block is defined. Another design variable for the H-Plane iris functional block might be the ratio of input port width to output port width. Mapping of this design variable to structural variables depends on external factors and must be accomplished through optimisation.

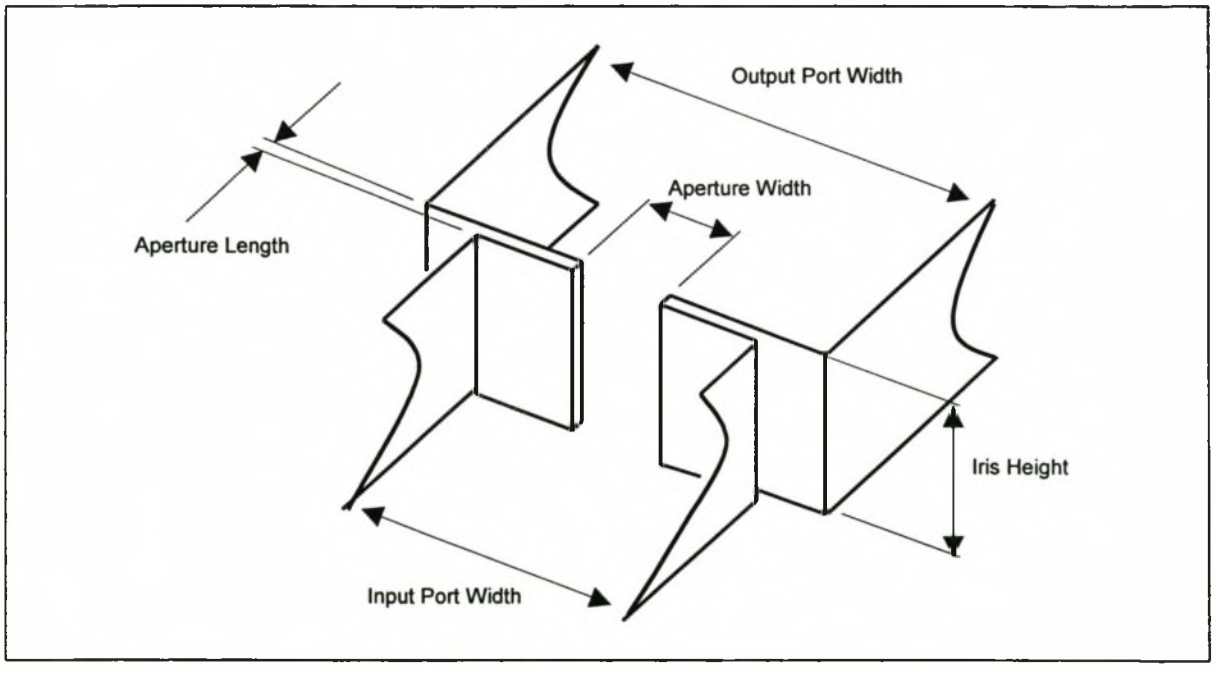


Figure 2.2 The structural variables of an H-Plane Iris

IMPLEMENTATION OF FUNCTIONAL BLOCKS

Functional blocks are implemented as individual software functions. Every functional block has to provide a number of services to various calling programs. These services are:

1) Block definition

Before design, the various structural variables of a functional block must be set to fixed or subject to design. These rules hold for every instance of the block that is used in the design, and are determined by the functional block itself through interaction with the user. At this point the user can also set limits for changeable variables if so desired. The output of this service is a list of structural rules and design variables that are used respectively by the Rubberband algorithm and by the design procedures.

2) Evaluation of design variables

This does the inverse mapping - from structural variables to design variables. The resulting list of design variables is used by the Rubberband algorithm.

3) Analysis

This service can do one of two things: it either returns the physical structure of the functional block to the program used to analyse the entire structure, or, if the block is self-analysing – implying that it provides its own electrical characteristics through some other means – it can return its own electrical characteristics depending on the current values of the structural variables.

A number of functional blocks were implemented for use in this investigation according to the criteria detailed above. A selection is shown in figure 2.3 with a list of their structural and design variables. Note that, in general, it was preferred not to specify the port dimensions directly as design variables, but rather to allow the ratios of the dimensions to be designable. The reason for this, is that this scheme allows the Rubberband algorithm more freedom to 'stretch' structures, incorporating new functional blocks more smoothly into the solution structure. This is not a prerequisite for the design strategies to work, but it was found that they generally tended to produce better results if this strategy were employed.

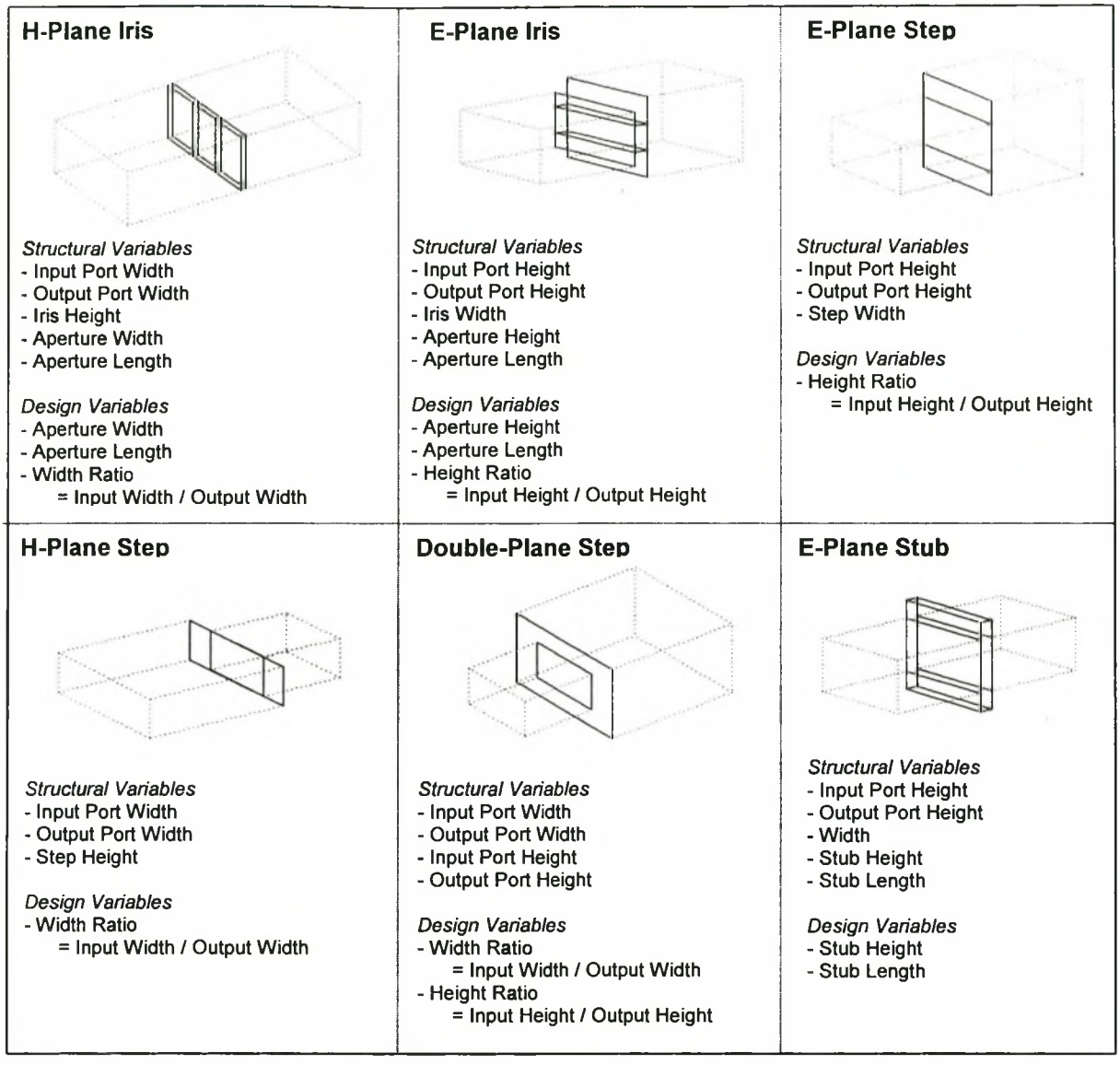


Figure 2.3 A selection of functional blocks and their variables

SELECTION OF FUNCTIONAL BLOCKS

Prior knowledge of the problem is best exploited through the process of selecting functional block types for a design. Although the conceptual design-level definition is tailored to allow the inclusion of a potentially large variety of different types of blocks in a design, the most efficient results are typically found by selecting only those types of functional blocks that are most likely to provide the excitation most needed for he specific problem. If all the blocks available were included, of which the above are only a selection, the design algorithms would be obliged to test each one's viability in the design. Although the wider range of possibilities might lead to a better solution, it would come at a disproportionately increased cost of design time.

Selection of blocks can be best accomplished by some knowledge of the properties of various types of blocks, which can be easily determined from numerical analysis. Chapter 3 presents a circuit model that can also be of great use in interpreting the behaviour of various structures on a modal level. The selection of appropriate functional blocks for the applications considered in this dissertation is motivated in the corresponding chapters that follow.

2.4.2 The Rubberband Algorithm

The Rubberband algorithm is responsible for transforming the conceptual design model, which is defined by design variables, to a physical structure that is to be evaluated. To do this, it must take into account the fixed structural variables, including the port specifications, and it must ensure that functional blocks can be appropriately connected to one another. This problem can become complex as certain functional blocks, such as the H-Plane iris, must show continuity in one or both port dimensions. Other functional blocks may have fixed port dimensions specified as at arbitrary points in the structure to which adjacent functional blocks must adhere. Depending on the fixed structural variable definitions, certain configurations of functional blocks may simply not be possible. If such an event occurs, the algorithm returns an error message to the design algorithm which typically abandons the unviable configuration.

In general, the assignment of values to structural variables that are not pinned down by other fixed structural variables or constraints is accomplished through optimisation using the least squares method. The error values are defined as the difference between the design variables reported by the functional blocks (due to the current structural variables) and the desired design variables specified by the design algorithm. The Rubberband algorithm minimizes the individual error values by optimising the free structural variables. In cases where the mapping from design variable to structural variable is trivial, this implies a simple variable assignment; in other cases, where design variable values may conflict with fixed parameters, the optimiser is employed to find a choice of free structural variables that best satisfies the design variables.

Figure 2.3 illustrates this for a simple cascade system of H-Plane waveguide steps. In this example, an original structure is specified with fixed port dimensions and known step width ratios as shown. The desired width ratio of the second step is then modified from its original

value of 0.42, whilst the other desired width ratios are specified to remain as they are in the original structure. The column on the left shows two structures that result from the Rubberband algorithm when the desired width ration of the second step is raised to 0.6 and 0.8. All the width ratios deviate form their desired values, shown in <>, so as to find a compromise structure that best suits all the design variables. The column on the right shows the effect of decreasing width ratio. In the last example, the algorithm has appeared to clip the widths as the dashed line. This is because the structure was constrained to a maximum size width in the constraint definitions.

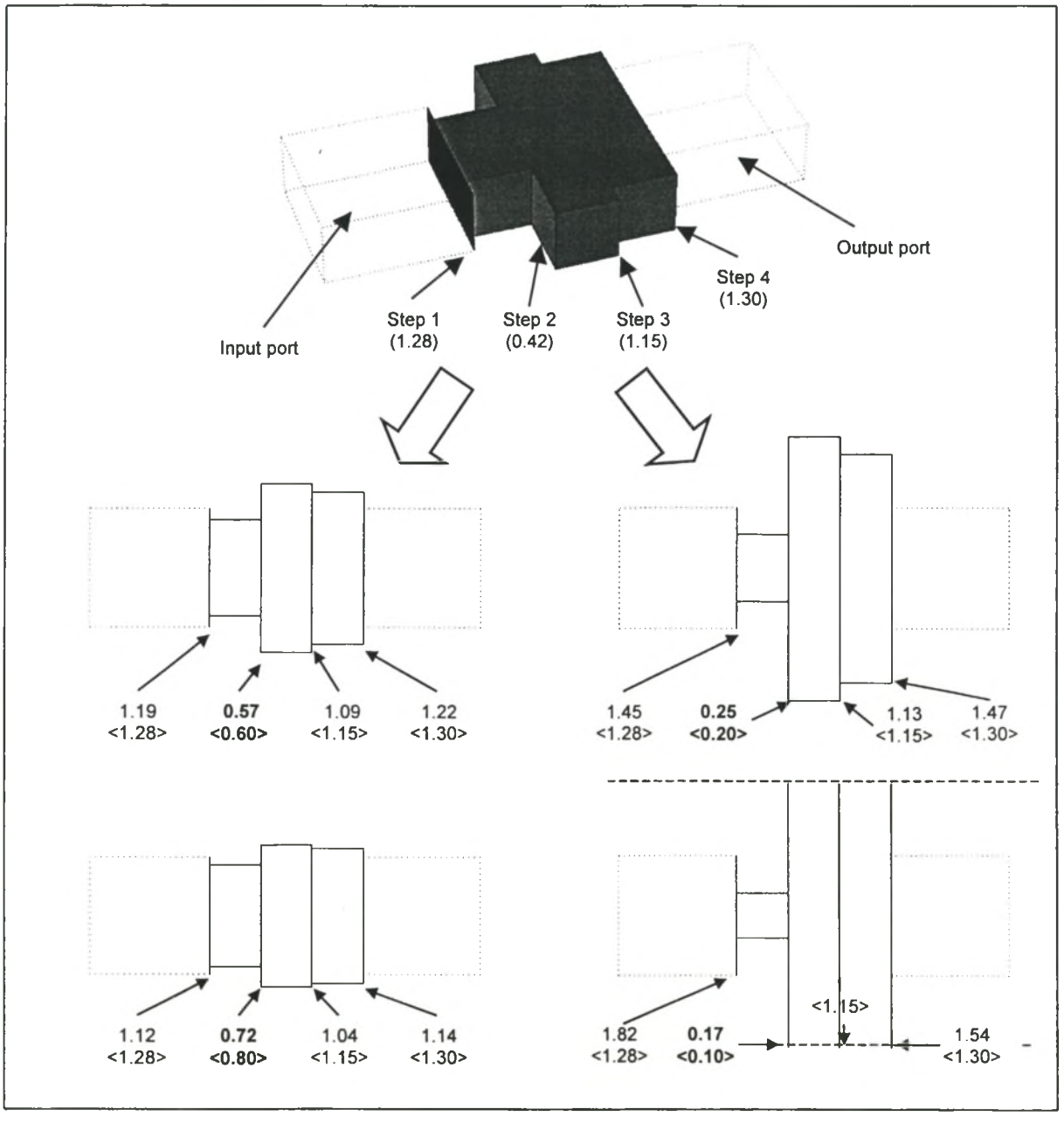


Figure 2.4 The Rubberband algorithm in action. (<> implies a desired width ratio)

The use of the Rubberband algorithm ensures that viable and well-conditioned physical structures emerge from the design variables. This allows the design procedures to operate with relative freedom and, most importantly, allows the seamless addition, subtraction, and merging of structures with ease – greatly facilitating the growth based approaches to be discussed in the rest of this chapter.

2.5 Growth-Based CAD Strategies

The general approach to be investigated is the gradual addition of functional blocks to one or more solution structures in such a way that each addition results in a significant improvement in the performance of the structure(s). Performance is to be evaluated according to the difference between the achieved and desired electrical response according to a sum of squares error function.

This section details a number of techniques that exploit this and any other information that might be available to accomplish design based on the general principle of growth of complexity of the solution. Two broad classes of design strategies are discussed here. The first pursues the addition of functional blocks in a systematic, scanning-based approach. The second uses a stochastic approach, based on an evolutionary strategy, to incorporate complexity in the structure gradually. Before these strategies are introduced, however, basic structural optimisation is first discussed.

2.5.1 Basic Structural Optimisation

The simplest application of computer aided design to developing solution structures is the use of direct optimisation of the various design variables. This investigation makes use of this simple strategy, not only in the refinement of solutions generated by the more sophisticated methods detailed below, but also as a component of some of these methods. It was not, however, found to serve as a viable design strategy in itself.

The application of direct optimisation in the context of functional block-based structural description is relatively simple. The user defines a 'design' error function, that evaluates the error corresponding to the difference between a desired and true electrical response. The true electrical response is obtained by analysing the structure that results from the application of the Rubberband algorithm to the set of design variables available to the optimizer. The design variables are made available by the functional blocks and line sections. The process is outlined in figure 2.5 below.

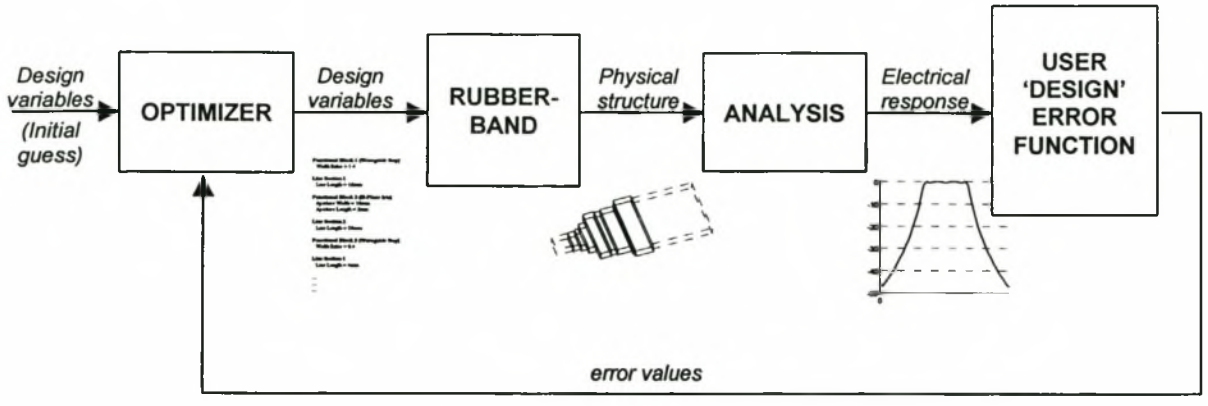


Figure 2.5 Basic structural optimisation

All structural optimisation used in this investigation is applied in this way and makes use of the hybrid least squares, steepest descent algorithm developed for this investigation and detailed in Appendix 1.

2.5.2 The Generalized Scanning Strategy

The scanning strategy systematically investigates the addition of the various available functional blocks at each viable insertion point along the existing structure. Insertion points correspond to transmission line sections of a length at least greater than a specified minimum and at points at the beginning and end of the structure. The basic strategy, detailed below, is additionally enhanced by a variety of modifications that are detailed separately in their own sections. All the CAD algorithms and their enhancements discussed here are presented more formally in Appendix 2.

THE BASIC SCANNING STRATEGY

This strategy requires, as a minimum, at least one error value relating to the fitness of the proposed solution structure. Its basic workings are explained here with the aid of a simple example. It is desired to produce a single mode bandpass filter from E-Plane steps in waveguide. The steps are modeled as impedance jumps and constitute the functional blocks used for this simple example design.

The algorithm starts by optimising a number of random seed structures, each comprising a minimum number of functional blocks that will satisfactorily realize a viable structure. This typically amounts to two functional blocks separated by a line section. The best outcome of these seeds is adopted as the starting structure. Figure 2.6 shows the random seeds before and after optimisation and their resultant electrical responses and sum squared errors.

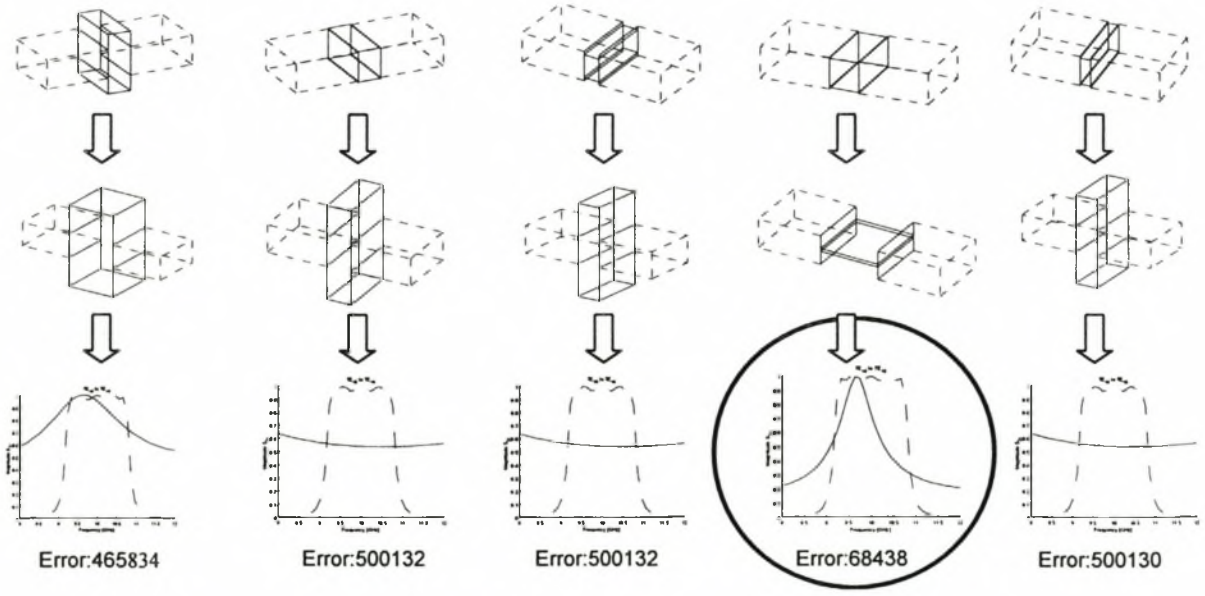


Figure 2.6 Optimisation of the random seeds and the best result

The algorithm then 'scans' the starting structure by attempting to place every available functional block, which, in this case, is limited to only E-plane steps, at each viable insertion point in the structure as shown in figure 2.7. For each successful insertion, two optimisations are carried out. The first is to optimise the design variables of the newly inserted functional block and the adjacent line lengths alone. In many cases this can be efficiently done by storing the electrical characteristics (e.g. S-parameters) of the unchanging structures on either side of

the newly inserted functional blocks, and by using those stored characteristics in the analysis rather than having to re-analyse the structure in its entirety. The example used here, however, does not readily allow this approach as the entire structure is affected due to the action of the Rubberband algorithm. The next optimisation is of all the available design variables in the new structure, including the design variables of the newly inserted block and line lengths.

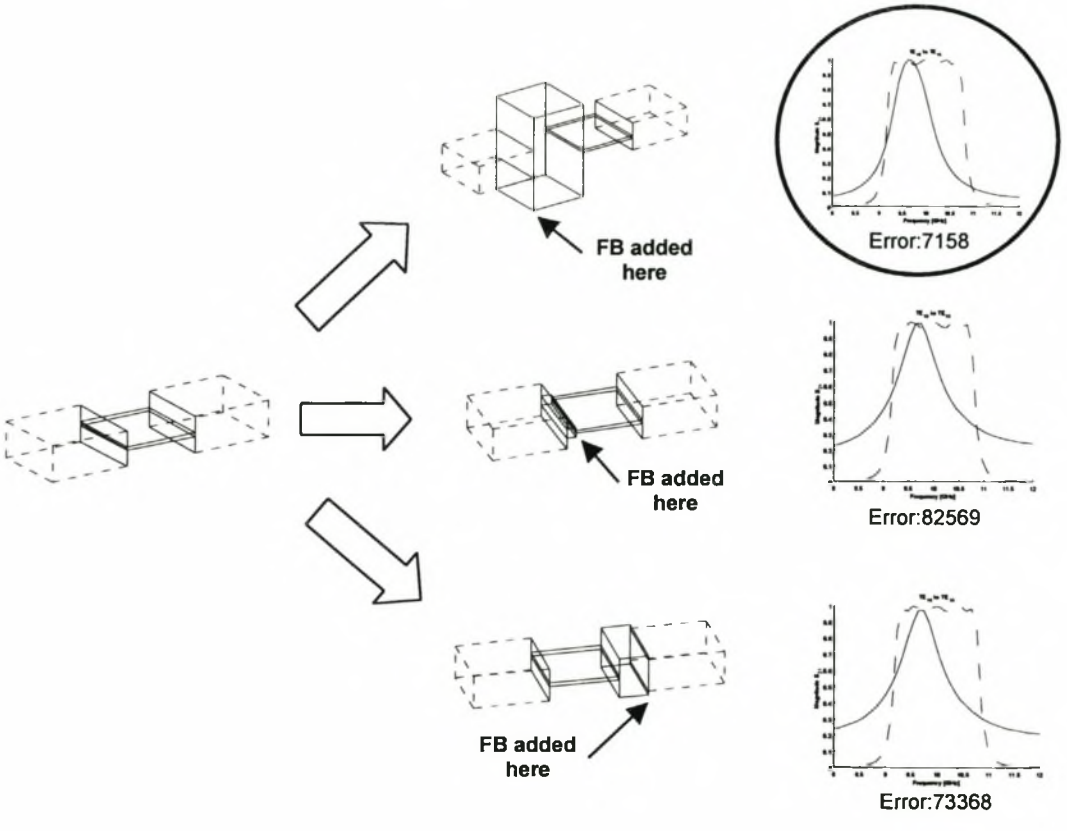


Figure 2.7 The first scan of the starting structure

Clearly the placement of a suitable functional block just behind the input port results in the best improvement in the structure's response. If this improvement is more than a specified minimum, the new functional block is allowed to be incorporated in the structure. Note that if the designer chooses to specify a limit on the structure's physical length, the algorithm will reject the insertion of any block that causes the new structure to exceed this specified maximum. Figure 2.8 shows the results of the next scan of the example design problem.

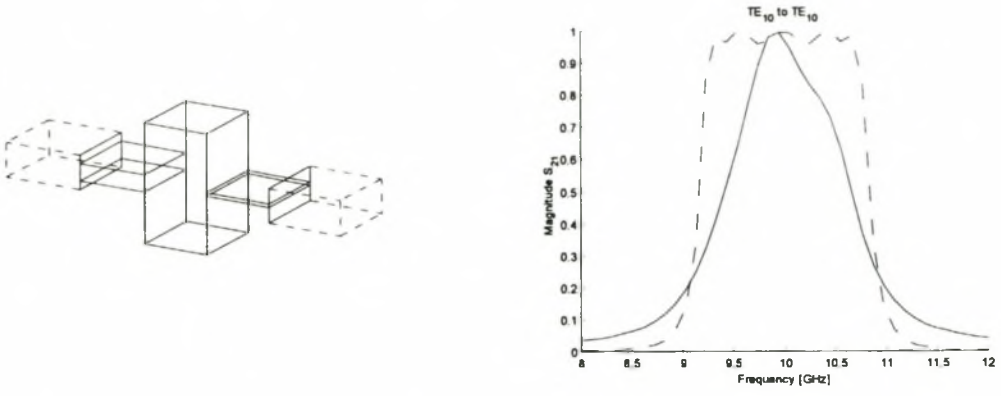


Figure 2.8 Result of next scan

This 'scanning' procedure is repeated over the new structure a number of times and the structure grows more complex until it either fulfils a specified sum squared error, or grows in complexity beyond a specified maximum. It is possible for a particular scan to fail to identify any new viable insertion. This is typically allowed to happen a limited number of times before it is concluded that no more improvement can be made and the algorithm terminates.

This strategy was found to work reliably for a number of applications, some of which are detailed in later chapters. It can, however, be improved in a number of ways, especially by the incorporation of information about the problem. The modifications and additions made to this basic strategy generally tend not to limit its flexible application, but to render a more elegant approach that guarantees more efficient and reliable results.

REMOVAL OF REDUNDANT SINGLE MODE LINE SECTIONS

In the example above, the maximum line length was set at 30mm by the Rubberband algorithm. This is about ¾ of a wavelength of the only mode under consideration, the TE_{10} mode. This was done because it is known that in a single mode environment line lengths of longer than ½ wavelength are (typically) unnecessary. In multimode problems, limiting the line length to such a low value manually is typically not desirable as it may exclude requirements for multimode sections of lines. Excessive line lengths can, however, still be avoided for line sections supporting only one propagating mode. Appendix 2 details the algorithm that can be included in the basic scan to check for this occurrence. The algorithm

effectively checks each line section after a successful scan, if a line section is found that supports only one propagating mode, which can be determined by the line's dimensions, and that line is longer than ½ wavelength of that mode at the center frequency of the design, then the algorithm shortens the line by ½ wavelength sections until it is less than ½ wavelength. It then re-optimises the structure. If the optimised shortened structure performs as well, or at least nearly as well as the pre-shortened structure, then it is adopted instead as the new current structure. If not, the algorithm reverts to using the original structure. This addition to the basic scan algorithm comes at very little cost, but has been found to be most useful in ensuring short solutions in some problems. A equivalent operation for multimode lines is potentially possible, but has not been implemented in this investigation.

REMOVAL OF REDUNDANT FUNCTIONAL BLOCKS

It can occur that a functional block, added early on in the design procedure, becomes defunct through the addition of a more suitable alternative later in the design. In such cases, the design variables of the block might be optimised in such a way as to make it completely unnecessary to the structure from an electrical point of view. In the case of, for example, an E-Plane iris, the aperture height might be set equal to the iris height. In cases where constraints do not allow this to happen, e.g. a specified maximum value of aperture height, the block might even be detrimental to the structure as a whole. Removal of the block would then typically provide improvements in the overall performance of the structure.

As before, this modification is applied directly after a successful scanning operation. The algorithm simply removes each functional block in turn and unites the adjacent line sections in a new random line length. After each removal, the simplified structure is re-optimised. It has been observed that in an overwhelming number of cases, the removal has a bad effect on the structure's performance, but in some cases, where the block is defunct, the removal might result in little or no degradation of the structure (and perhaps even in its improvement) and the simplified structure can be adopted as the new current structure. The ability to remove functional blocks of this nature provides a way of ensuing the least possible complexity for a specific performance. This can help greatly in speeding up the design, analysis and fabrication of the structure, but these benefits can be surpassed by the added extra optimisations required to implement the modification.

LINE LENGTH GRID SEARCH OR USE OF PRIOR KNOWLEDGE

A weakness of the basic scanning technique is the degree of randomness employed in the design. This is especially so in the selection of the line lengths separating the newly-inserted functional block and the rest of the structure. Experimentation with the design of narrow band single mode filters has revealed that it can be difficult for the first optimisation step to acquire the best possible separating line lengths unless good fortune led to especially suitable random starting values. This is because design for resonance results in a very badly conditioned error function. Figure 2.9 plots the sum squared error of a typical error function for a single mode narrow band filter as the length separating two of the functional blocks (H-Plane Irises) is varied. Not only is the curve convex far from the optimal point, but there are two local minima adjacent to the optimal point that threaten to snag the optimiser, effectively stopping it from reaching a successful design.

Although it might be possible to reformulate the error function to avoid these undesirable properties, it would not address the problem with the design strategy. An alternative approach uses a constrained grid search to select the new line lengths before the first optimisation step. Again, the cost is increased effort in each scan, but specifically for narrow band designs this pays off with much more reliable scans. A generalization of this approach is to make use of any information about the problem to select the appropriate starting values for the line sections. This can greatly increase the reliability of the individual scans, resulting in a fast design process.

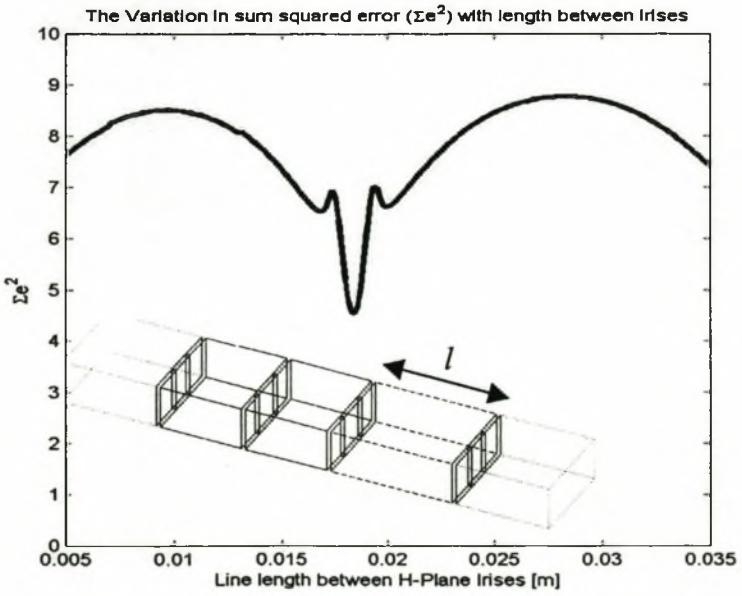


Figure 2.9 The problems associated with optimising line lengths

THE RECURSIVE TREE SEARCH STRATEGY

A generalization of the basic scanning strategy uses a tree search implementation. Instead of dealing with a single structure at a time by selecting the best solution for every scan, this approach retains a number of viable, growing solutions and attempts to improve each one in turn. The implementation of this approach simply renders the scanning technique into a generalized recursive algorithm that calls itself at the end of a scan, for each successful outcome of the scan. This effectively implies that no improved structure is discarded – resulting in a much higher reliability.

The cost, however, is very high, as the number of solutions under investigation grows exponentially if not curtailed somehow. For this reason, direct implementation of the tree search generalization was not attempted in this investigation. A simplified implementation was, however, developed that proved to be a great use, as will be shown in a later chapter. This implementation does not allow a complete scan. Rather, it limits insertion of a new functional block only at the end of the structure, reducing the associated costs of optimisation and analysis to a level that makes the magnitude of the tree search computationally viable. It can, therefore, be seen as a special case of the general strategy outline above. This simplified tree search strategy algorithm is detailed in the appendices.

Initialization of the algorithm is similar to the basic scanning technique, with a number of starting seeds optimised in the usual way. However, instead of selecting only the best solution, a proportional value of the resultant structures is retained. For each structure in this set of solutions, a new instance of the algorithm is called. When the algorithm receives a structure to work on, instead of scanning the entire structure, it investigates insertion of all the available blocks only at the end of the structure, applying either optimisation or prior knowledge with respect to the incorporation of the functional block into the structure as usual. Typically, a number of insertions will be found to improve the structure's performance. Of these, a proportional value are selected to start a new branch of the tree search. If no insertions are found to improve the performance of the structure adequately, the branch is killed off and the search curtailed. Figure 2.10 schematically illustrates the growth of the search from the starting seeds.

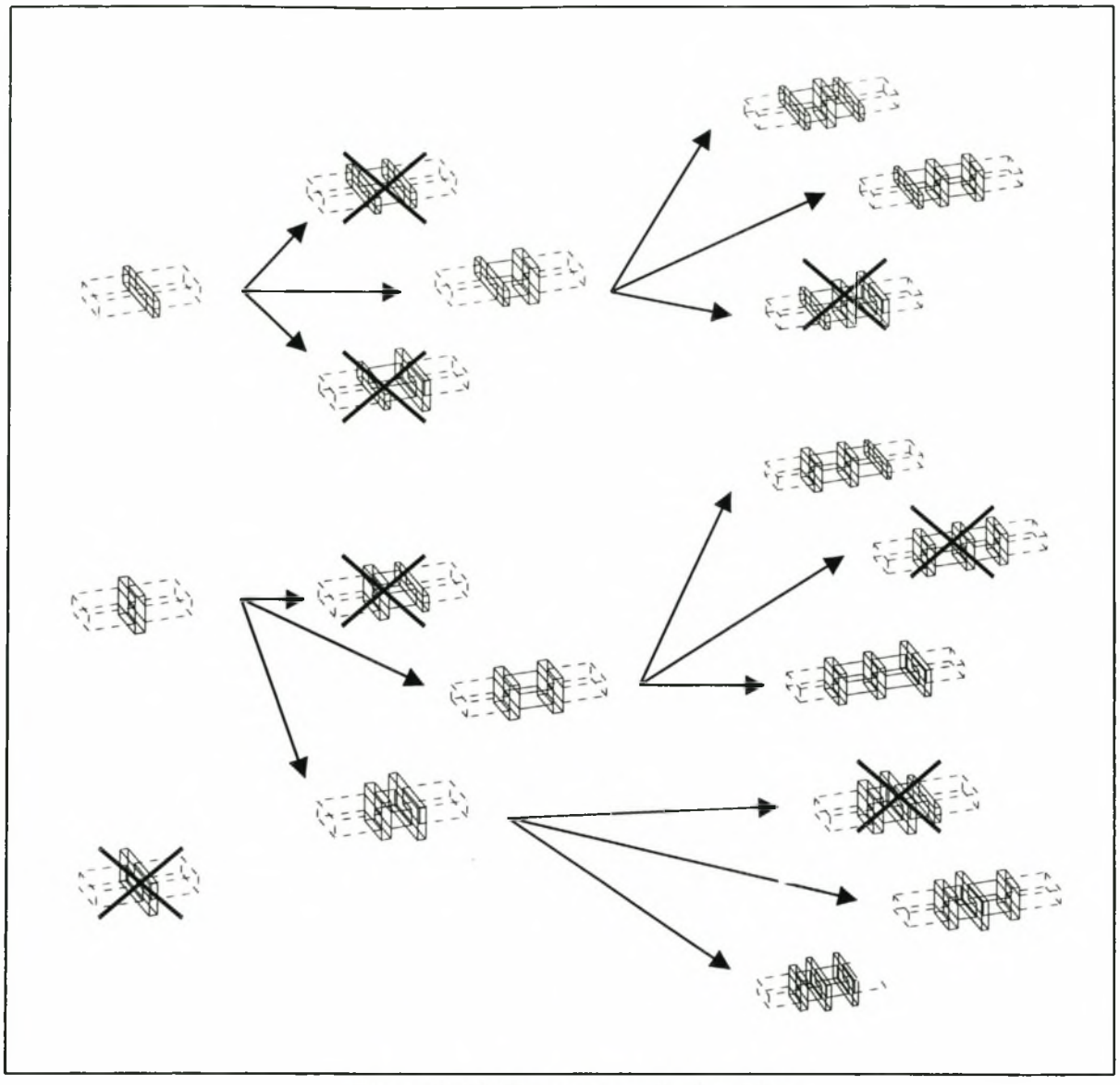


Figure 2.10 The tree search algorithm

Note that termination of the search may occur 'naturally' as the structures reach a level where they simply cannot be improved by the further addition of a functional block. Alternatively, an artificial limit on the growth of the solutions may be specified, such as maximum structural length or complexity. An interesting alternative that was found to produce good results was to limit the number of times a functional block can be used in the structure – thereby producing a search that finds the best configuration of a limited number of blocks of various types in a solution structure.

The generalized tree search implementation of the scanning strategy searches a much larger area of design space, suggesting that it may provide more reliable design solutions that are

'globally optimal'. The cost of this approach is computational effort which makes it difficult to evaluate given the current level of technology. The method is, however, well suited to parallelization, and its application in this respect might well be an area of fruitful research. The pursuit of improved reliability with the aid of a more extensive list of concurrently growing solutions can, however, alternatively be accomplished with the aid of evolutionary strategies.

2.5.3 The Evolutionary Strategy

This strategy makes use of the principle of evolution to guide the growth and development of large quantities (populations) of candidate structures concurrently. The use of evolutionary strategies for circuit design and optimisation is not new. It has been successfully used in a variety of applications, including the design of analog circuits [Koza 1997]. An excellent summary of the typical uses and implementations of various evolutionary techniques, including genetic algorithms, is due to Bäck et al. [Bäck 1997]. Tang et al. also provide an excellent outline of the general structure of genetic algorithms and their uses [Tang 1996]. The great strength of the evolutionary approach is great flexibility and excellent global minima seeking properties, as has been widely recognized in the literature. This makes it eminently suited to the problem inherent to multimode design.

The evolutionary strategy developed here is specifically intended to facilitate the growth of the population's complexity gradually as the population evolves, thereby mimicking nature more closely than standard evolutionary approaches. It also makes use of the great flexibility of the functional block-based description of waveguide structures. The algorithm implemented for this investigation is detailed more thoroughly in the appendices.

The algorithm starts by initializing a population of random simple one- or two- functional block structures (organisms) similar to the seeds created in the scanning strategies. This population is operated on by the genetic methods of competition and reproduction to gradually improve performance of the population by evolving more complex structures. Competition is accomplished with the tournament method, which pairs random structures and chooses the one with the lowest sum squared error. If the difference in sum squared error

between the two candidates is smaller than a specified percentage, then the structure with smaller physical length (as reported by the Rubberband algorithm) is instead favored. This encourages the selection of physically short structures. The 50% of the population that is discarded is replaced by 'offspring' of the fitter 50% of the population.

Reproduction is in one of two modes. Asexual reproduction operates on a single random member of the parent population and is capable of modifying a connecting line length or of adding, subtracting, moving, or modifying a functional block in the parent structure. Sexual reproduction, which involves two random parents, splits both parents at a random point in their structure and combines two of the halves to form a new structure. The early part of the design process favors asexual reproduction, but as the solution structures increase in complexity, sexual reproduction is favored.

This strategy incorporates a rather large degree of 'elitism' in the retention of 50% of the existing generation for the next generation. It is, however, important to note that the expense of analysis lies only with the newly reproduced 50% of the population. The fitness of the older 50% has already been established and need not be re-computed.

The implementation described above is by no means the only viable method of exploiting evolutionary strategies for this purpose. Many other approaches exist that use alternative methods of competition and selection for parenthood. These would doubtless leads to results similar to or perhaps even better than those obtained by the method outline here. The essential component of this approach is the ability of the evolutionary strategy to control the evolution of the population's complexity by starting with a simple population and favoring asexual reproduction early on in the design process.

The performance of the evolutionary strategy in a practical example is detailed in chapter 5.

2.6 Conclusion

This chapter has proposed a growth-based design strategy for the solution of difficult design problems in multimode waveguide. The implementation of these design strategies is enhanced by the development of a conceptual functional block-based framework for the description of arbitrary structures in waveguide. This framework allows the easy addition of discontinuities into an existing structure and is designed to minimize the effect such inclusions will have on the rest of the physical structure.

The functional blocks are controlled by a Rubberband algorithm, which ensures that the physical structure represented by the arrangement of functional blocks obeys physical constraints imposed on the system.

Two classes of growth-based strategies have been introduced and thoroughly explained. The first class uses a scanning-based approach to systematically insert new functional blocks into a growing solution structure. The second makes use of an evolutionary strategy to allow a population of structures to evolve from simple single-functional block structures to more complex structures with more favorable electrical characteristics.

Chapter 3

Analysis Technique Development

3.1 Introduction

The development of efficient analysis techniques is crucial to effective CAD. In recent years, rapid developments in numerical analysis theory, as well as increasing computational capacity, have greatly enhanced the possibilities for CAD in the domain of passive microwave components. The trade-off between speed, accuracy and flexibility, however, still remains a problem. Until general purpose techniques, like FEM and FDTD, develop to the extent that they can compete in terms of speed and accuracy with more specialized techniques, like MoM and Mode Matching, the designer will be forced to limit the freedom of form that his designs may assume by opting for a more specialized analysis technique. Since, in the context of this dissertation, only the behaviour of multimode devices made up of cascaded double-plane steps in waveguide is considered, the specialized analysis method that emerges as most favourable is the mode matching method.

This chapter will start by discussing the mode matching formulation for double-plane steps in waveguide. Next, techniques for cascading multimode discontinuities in waveguide will be assessed and a suitable method will be adopted for use. By this stage, full analysis of prototype structures will be possible. The following sections will present a number of original developments that greatly enhance the mode matching formulation:

- An equivalent circuit formulation for the mode matching operation.
- An intelligent strategy for mode selection that allows very fast analysis of cascaded step structures.
- The use of surrogate models with the aid of a basic interpolation technique that will be used approximate EM analyses of functional blocks.
- Modal transforms that allow conversion of network parameters from one modal type (e.g. TE-to-x/TM-to-x) to another.

3.2 The Mode Matching Method in Rectangular Waveguide

The mode matching method was first proposed by Alvin Wexler in 1967 [Wexler 1967]. The usefulness and accuracy of his method has assured it a great deal of attention from researchers [Omar 1985, Sorrentino 1991, Alessandri 1992, Eleftheraides 1994]. In its present form, mode matching is used extensively for the analysis of many types of waveguide and planar circuit discontinuity problems and promises an exciting future through its potential hybridization with other codes [Meyer 1995].

Mode matching entails the expansion of tangential fields on both sides of a discontinuity into infinite summations of modes with unknown complex amplitudes. Field amplitudes are determined by enforcing boundary conditions at the discontinuity plane and by using the principle of orthogonality to extract the individual modal contributions. In order to render the technique numerically viable, only a finite number of modes in each guide must be considered. This creates a convergence issue that has been discussed in the literature [Leroy 1983]. The essence of this issue is that a sufficient number of modes must be used in the correct ratio to ensure convergence. For discontinuities that differ markedly from the embedding structure, this number can be very high.

3.2.1 Theoretical Development

The development presented here combines the approaches found in various sources [Shih 1989, Alessandri 1992]. The formulation is capable of analysing general steps in rectangular waveguide with perfectly conducting walls as shown in figure 3.1. Results are presented in terms of the complete normalized scattering parameters for both sets of modes on either side of the step discontinuity.

a) Modal description of fields

In the case of air-filled rectangular waveguide with perfectly conducting walls, wave propagation can described by the Helmholtz equations for phasor fields as shown in eqn. (3.1) [Ramo 1993a].

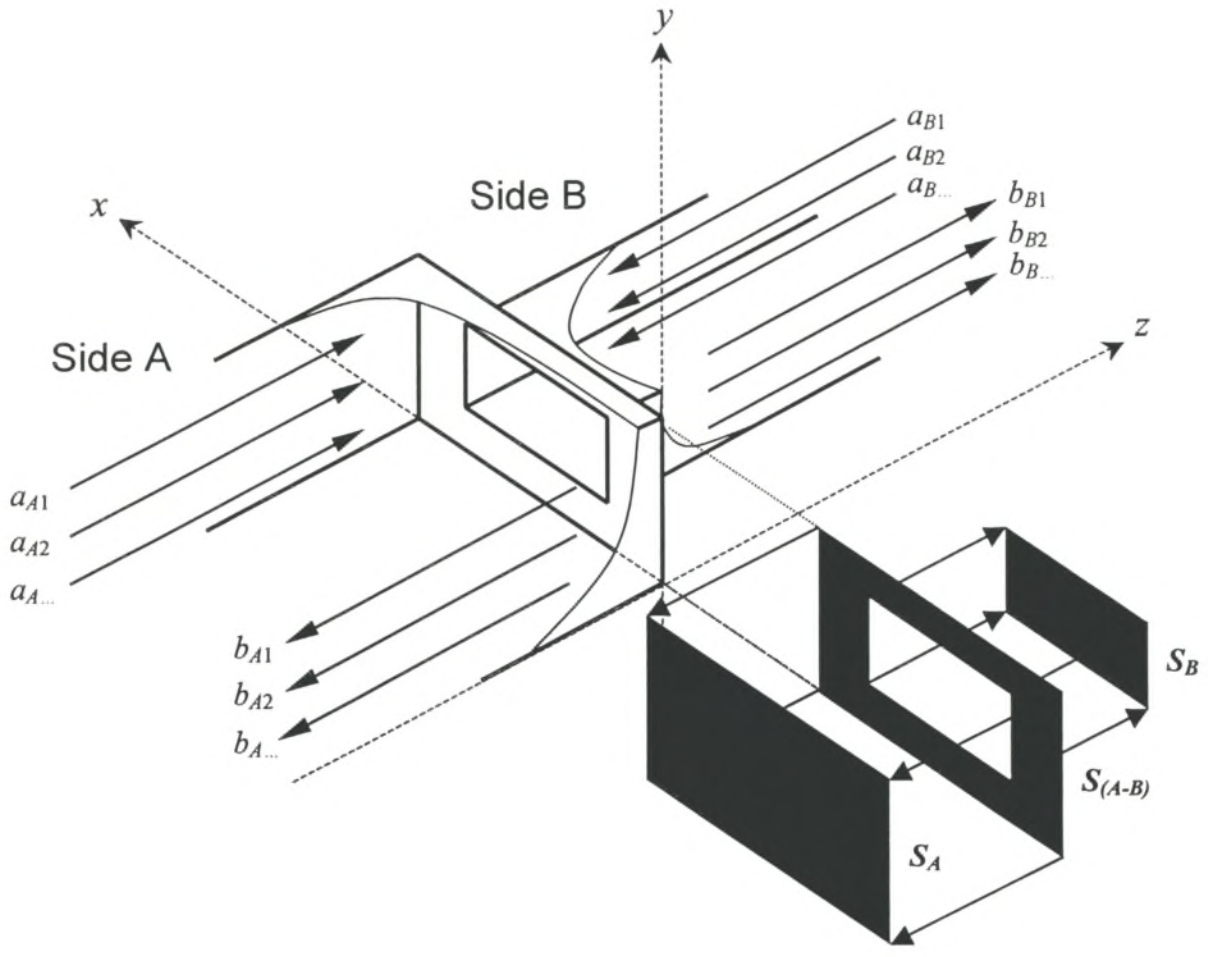


Figure 3.1 A Double-Plane Step in Rectangular Waveguide

$$\begin{aligned} \nabla^2\mathbf{E} &= -k^2\mathbf{E} \\ \nabla^2\mathbf{H} &= -k^2\mathbf{H} \\ k^2 &= \omega^2\mu_0\varepsilon_0 \end{aligned} \tag{3.1}$$

These can be solved in terms of the propagating **E** and **H** fields in a number of linearly dependant solutions known as the modal solutions. Of the infinitum of available solution sets, three standard types can be identified and classified according to field behaviour. They are the TE_x/TM_x (TE-to-x and TM-to-x) modes, the TE_y/TM_y modes, and the most widely used TE_z/TM_z modes. The mode matching formulation may be expressed in terms of any one of these three modal solutions, with the scattering parameter results relating only to the chosen modal field representation.

For the purposes of this investigation the most generally applicable modal set, the TE_z/TM_z solution is chosen for the mode matching formulation. Transforms to the other modal sets are presented in section 3.7.

The field magnitudes for both the TE_z and TM_z modes are well known [Ramo 1993b] and are summarized in eqns. (3.2) and eqns. (3.3), with C_{TM} and C_{TE} arbitrary constants. The propagation constant, γ, is either purely imaginary (propagating modes) or purely real (non-propagating modes). a and b are the waveguide dimensions in the x- and y-directions respectively. m and n are the mode indices, neither of which can be zero for the TM modes.

TM Fields

$$E_x = C_{TM}\left[-\frac{\gamma k_x}{k_c^2}\cos(k_x x)\sin(k_y y)\right]$$
$$E_y = C_{TM}\left[-\frac{\gamma k_y}{k_c^2}\sin(k_x x)\cos(k_y y)\right]$$
$$E_z = C_{TM}\left[\sin(k_x x)\cos(k_y y)\right]$$
$$H_x = C_{TM}\left[\frac{j\omega\varepsilon k_y}{k_c^2}\sin(k_x x)\cos(k_y y)\right]$$
$$H_y = C_{TM}\left[-\frac{j\omega\varepsilon k_x}{k_c^2}\cos(k_x x)\sin(k_y y)\right]$$
$$H_z = 0$$

TE Fields

$$E_x = C_{TE}\left[\frac{j\omega\mu k_y}{k_c^2}\cos(k_x x)\sin(k_y y)\right]$$
$$E_y = C_{TE}\left[-\frac{j\omega\mu k_x}{k_c^2}\sin(k_x x)\cos(k_y y)\right]$$
$$E_z = 0$$
$$H_x = C_{TE}\left[\frac{\gamma k_x}{k_c^2}\sin(k_x x)\cos(k_y y)\right]$$
$$H_y = C_{TE}\left[\frac{\gamma k_y}{k_c^2}\cos(k_x x)\sin(k_y y)\right]$$
$$H_z = C_{TE}\left[\cos(k_x x)\sin(k_y y)\right]$$

(3.2)

Wave Number and Propagation Constant

$$\gamma^2 = k_c^2 - k^2 \qquad k_x = \frac{m\pi}{a}$$
$$k_c^2 = k_x^2 + k_y^2 \qquad k_y = \frac{n\pi}{b}$$

(3.3)

b) Normalization of the modal field magnitudes for unity power magnitude

It is essential to have the resultant s-parameters normalized to maintain numerical uniformity and to attach a more meaning physical value to the parameters. Values for the constants, C_{TM} and C_{TE}, for each mode are therefore chosen such that the magnitude of the complex power over the z-directed surface, S_{WG}, given in eqn. (3.4), is unity.

$$\mathrm{P} = \iint_{S_{WG}} \left[\left(\mathbf{E} \times \mathbf{H}^* \right) \cdot \hat{i}_z \right] \cdot dA \tag{3.4}$$

The phases of C_{TE} and C_{TM}, will have no effect on the phase of the complex power, which is always ±90 degrees if the mode is not propagating and zero degrees if the mode is propagating. The choice of the phase of these constants is therefore arbitrary, but it will affect the angles of the off diagonals of the generalized scattering matrix. Provided the same normalizing convention is maintained, this will have no effect on the magnitudes of the scattering parameters of cascaded systems. Appendix 3 derives the normalization constants and motivates the choice of their phase. The tangential components of the normalized TE and TM modes are given in vector notation in eqns. (3.5) and (3.6) respectively with the modal wave impedances and admittances as in eqns. (3.7) and (3.8).

$$\begin{aligned} \mathbf{E}'_{TE} &= \sqrt{Z_{TE}} \cdot \frac{q}{\sqrt{ab}} \cdot \frac{1}{k_c} \left[k_y \cos(k_x x) \sin(k_y y) \cdot \hat{i}_x - k_x \sin(k_x x) \cos(k_y y) \cdot \hat{i}_y \right] = \sqrt{Z_{TE}} \cdot \tilde{e}_{TE} \\ \mathbf{H}'_{TE} &= \sqrt{Y_{TE}} \cdot \frac{q}{\sqrt{ab}} \cdot \frac{1}{k_c} \left[k_x \sin(k_x x) \cos(k_y y) \cdot \hat{i}_x + k_y \cos(k_x x) \sin(k_y y) \cdot \hat{i}_y \right] = \sqrt{Y_{TE}} \cdot \tilde{h}_{TE} \end{aligned} \tag{3.5}$$

$$\begin{aligned} \mathbf{E}'_{TM} &= \sqrt{Z_{TM}} \cdot \frac{q}{\sqrt{ab}} \cdot \frac{1}{k_c} \left[-k_x \cos(k_x x) \sin(k_y y) \cdot \hat{i}_x - k_y \sin(k_x x) \cos(k_y y) \cdot \hat{i}_y \right] = \sqrt{Z_{TM}} \cdot \tilde{e}_{TM} \\ \mathbf{H}'_{TM} &= \sqrt{Y_{TM}} \cdot \frac{q}{\sqrt{ab}} \cdot \frac{1}{k_c} \left[k_y \sin(k_x x) \cos(k_y y) \cdot \hat{i}_x - k_x \cos(k_x x) \sin(k_y y) \cdot \hat{i}_y \right] = \sqrt{Y_{TM}} \cdot \tilde{h}_{TM} \end{aligned} \tag{3.6}$$

$$\mathrm{q} = \sqrt{2} \times \mathrm{sign}(m) + \sqrt{2} \times \mathrm{sign}(n)$$

$$Z_{TE} = \frac{1}{Y_{TE}} = \frac{j\omega\mu}{\gamma} \tag{3.7}$$

$$Z_{TM} = \frac{1}{Y_{TM}} = \frac{\gamma}{j\omega\varepsilon} \tag{3.8}$$

Note the general properties of the modified field patterns, $\tilde{e}_i$ and $\tilde{h}_j$, due to orthogonality in eqn. (3.9).

$$\iint_{S_{WG}} \left[(\tilde{e}_i \times \tilde{h}_j^*) \cdot \hat{i}_z \right] \cdot dA = \iint_{S_{WG}} \left[(\tilde{e}_i \times \tilde{h}_j) \cdot \hat{i}_z \right] \cdot dA = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{3.9}$$

c) Waveguide fields expressed as a sum of the modes

If all the modes that can exist on either side of the discontinuity are arranged in order of increasing k_c, and indexed 1..N on side A and 1..M on side B, and the complex amplitudes of the modes are denoted a_A and a_B for modes travelling towards the discontinuity on sides A and B respectively, and b_A and b_B for modes travelling away from the discontinuity on sides A and B respectively (see figure 3.1), then an arbitrary tangential field pattern on side A can be expressed as in eqns. (3.10) and on side B as in eqns. (3.11)

$$\mathbf{E}'_A(z) = \sum_{\substack{k=1 \\ N\to\infty}}^{N} \left(a_{Ak}e^{-\gamma_k z} + b_{Ak}e^{+\gamma_k z}\right)\cdot \mathbf{E}'_{Ak} \qquad \mathbf{H}'_A(z) = \sum_{\substack{k=1 \\ N\to\infty}}^{N} \left(a_{Ak}e^{-\gamma_k z} - b_{Ak}e^{+\gamma_k z}\right)\cdot \mathbf{H}'_{Ak} \tag{3.10}$$

$$\mathbf{E}'_B(z) = \sum_{\substack{p=1 \\ M\to\infty}}^{M} \left(a_{Bp}e^{+\gamma_p z} + b_{Bp}e^{-\gamma_p z}\right)\cdot \mathbf{E}'_{Bp} \qquad \mathbf{H}'_B(z) = \sum_{\substack{p=1 \\ M\to\infty}}^{M} \left(-a_{Bp}e^{+\gamma_p z} + b_{Bp}e^{-\gamma_p z}\right)\cdot \mathbf{H}'_{Bp} \tag{3.11}$$

The sign change evident in the expression for the B-side H field is due to the positive z-directed orientation of the Poynting vector of the field patterns, $\mathbf{E}'_{Bp}$ and $\mathbf{H}'_{Bp}$. Equivalent voltages and currents can now be defined in terms of the complex modal amplitudes as in eqn. (3.12).

$$\begin{aligned} V_{\{Ak,Bp\}} &= \sqrt{Z_{\{Ak,Bp\}}}\left(a_{\{Ak,Bp\}} + b_{\{Ak,Bp\}}\right) \\ I_{\{Ak,Bp\}} &= \sqrt{Y_{\{Ak,Bp\}}}\left(a_{\{Ak,Bp\}} - b_{\{Ak,Bp\}}\right) \end{aligned} \tag{3.12}$$

The Z's and Y's in eqn. (3.12) correspond to modal wave impedances and admittances as they appear in eqns. (3.7) and (3.8). The field distributions can be expressed in terms of these equivalents as in eqns. (3.13) and (3.14). Since mode matching only concerns itself with the fields at the discontinuity (z = 0), the exponent terms appearing in eqns. (3.10) and (3.11) can be neglected in the rest of the development.

$$\mathbf{E}'_A = \sum_{\substack{k=1 \\ N\to\infty}}^{N} V_{Ak}\cdot \tilde{e}_{Ak} \qquad \mathbf{H}'_A = \sum_{\substack{k=1 \\ N\to\infty}}^{N} I_{Ak}\cdot \tilde{h}_{Ak} \tag{3.13}$$

$$\mathbf{E}'_B = \sum_{\substack{p=1 \\ M\to\infty}}^{M} V_{Bp}\cdot \tilde{e}_{Bp} \qquad \mathbf{H}'_B = \sum_{\substack{p=1 \\ M\to\infty}}^{M} -I_{Bp}\cdot \tilde{h}_{Bp} \tag{3.14}$$

d) Enforcing boundary conditions

Boundary conditions have to be satisfied for the tangential fields at position z = 0 in both waveguides. Referring to figure 3.1, S_A and S_B are defined as the interior cross-sections of side A and side B respectively. $S_{(A-B)}$ is the conducting surface comprising the step between the two waveguides. The boundary conditions can be listed as follows and arise owing to a lack of surface current and invariant media on S_B.

Tangential **E** field of side A equal to zero on $S_{(A-B)}$. (Conductor)

$$\sum_{\substack{k=1 \\ N\to\infty}}^{N} V_{Ak} \cdot \tilde{e}_{Ak} = 0 \qquad \text{over } S_{(A-B)} \tag{3.15}$$

Tangential **E** field of side A equal to tangential **E** field of side B on S_B.

$$\sum_{\substack{k=1 \\ N\to\infty}}^{N} V_{Ak} \cdot \tilde{e}_{Ak} = \sum_{\substack{p=1 \\ M\to\infty}}^{M} V_{Bp} \cdot \tilde{e}_{Bp} \qquad \text{over } S_B \tag{3.16}$$

Tangential **H** field of side A equal to tangential **H** field of side B on S_B.

$$\sum_{\substack{k=1 \\ N\to\infty}}^{N} I_{Ak} \cdot \tilde{h}_{Ak} = \sum_{\substack{p=1 \\ M\to\infty}}^{M} -I_{Bp} \cdot \tilde{h}_{Bp} \qquad \text{over } S_B \tag{3.17}$$

e) Limiting the Number of modes

This step imposes a finite limit to the number of modes used in the summations of equations (3.15-3.17). This moves the development from that of theoretical field theory to practical numerical analysis. The assumption is made that the details of the extremely high order modes (those with high values of k_c), are of little interest, as excitation is primarily with lower order modes. As such, it is expected that the associated extremely high order modes will have very small magnitudes and can therefore be neglected, resulting in residual error in the complex amplitudes of the lower order modes. The strategy of dropping high order modes is generally accepted as reasonable practice in the literature [Eleftheraides 1994]. Section 3.5 will introduce an alternative strategy that provides a more intelligent estimate of which modes to

neglect. It is, however, generally observed that accuracy and computational effort increase with the number of modes taken into account on both sides of the discontinuity.

Of some importance is the relationship between the number of modes retained on either side. Some sources [Alessandri 1988, Leroy 1983, Shih 1989] indicate that the ratio of the number of modes on side A to the number of modes on side B should be the same as the ratio of their respective areas in order to maximize accuracy. [Accatino 1994] indicates that the modes on both sides should share similar values of k_c, which amounts to very much the same as the area criterion. It has been observed that using the same number of modes on both sides for widely varying cross-sectional areas can lead to convergence to an incorrect value. In this formulation N and M are constrained to adequately large integers with a relationship that obeys the area criterion.

f) The Mode Matching Equations

Limiting M and N to finite values, whilst enforcing tangential E-field equality allows eqn. (3.16) to be rewritten as eqn. (3.18). For an arbitrary A-side mode, q, eqn. (3.19) then results.

$$\sum_{k=1}^{N} V_{Ak} \cdot \tilde{e}_{Ak} = \sum_{p=1}^{M} V_{Bp} \cdot \tilde{e}_{Bp} \qquad \text{(over } S_B\text{)} \tag{3.18}$$

$$\therefore \sum_{k=1}^{N} V_{Ak} \iint_{S_B} \left[\tilde{e}_{Ak} \times \tilde{h}_{Aq} \cdot \hat{i}_z \right] \cdot dA = \sum_{p=1}^{M} V_{Bp} \iint_{S_B} \left[\tilde{e}_{Bp} \times \tilde{h}_{Aq} \cdot \hat{i}_z \right] \cdot dA \tag{3.19}$$

The orthogonal properties of the modified field patterns (eqn. (3.9)), combined with the first boundary condition can now be exploited in eqns (3.20), leading to an expression for the q-th equivalent voltage in eqn (3.21).

$$\begin{aligned} \iint_{S_B} \left[\tilde{e}_{Ak} \times \tilde{h}_{Aq} \cdot \hat{i}_z \right] \cdot dA &= \iint_{S_A} \left[\tilde{e}_{Ak} \times \tilde{h}_{Aq} \cdot \hat{i}_z \right] \cdot dA \qquad \text{(First Boundary Condition)} \\ &= \begin{cases} 0, & k \neq q \\ 1, & k = q \end{cases} \qquad \text{(By orthogonality)} \end{aligned} \tag{3.20}$$

$$V_{Aq} = \sum_{p=1}^{M} V_{Bp} \iint_{S_B} \left[\tilde{e}_{Bp} \times \tilde{h}_{Aq} \cdot \bar{z} \right] \cdot dA \tag{3.21}$$

If q is now allowed to assume values from 1 to M, eqn. (3.21) can be expressed more conveniently matrix notation in eqn. (3.22).

$$\bar{V}_A = [W]\bar{V}_B$$
$$\text{with } W(k,p) = \iint_{S_B}\left[\tilde{e}_{Bp} \times \tilde{h}_{Ak} \cdot \hat{i}_z\right] \cdot dA \tag{3.22}$$

In a similar way, for equality of tangential H-fields, eqns (3.23), (3.24) and (3.25) result when considering an arbitrary B-side mode, r.

$$\sum_{k=1}^{N} I_{Ak} \cdot \tilde{h}_{Ak} = \sum_{p=1}^{M} -I_{Bp} \cdot \tilde{h}_{Bp} \qquad (\text{over } S_B) \tag{3.23}$$

$$\therefore \sum_{k=1}^{N} I_{Ak} \iint_{S_B}\left[\tilde{e}_{Br} \times \tilde{h}_{Ak} \cdot \hat{i}_z\right] \cdot dA = \sum_{p=1}^{M} -I_{Bp} \iint_{S_B}\left[\tilde{e}_{Br} \times \tilde{h}_{Bp} \cdot \hat{i}_z\right] \cdot dA \tag{3.24}$$

$$\sum_{k=1}^{N} I_{Ak} \iint_{S_B}\left[\tilde{e}_{Br} \times \tilde{h}_{Ak} \cdot \hat{i}_z\right] \cdot dA = -I_{Br} \tag{3.25}$$

Allowing r to assume values from 1 to M and expressing the result in matrix notation leads to eqn. (3.26) with the $[W]$ matrix defined as in eqns. (3.22).

$$\bar{I}_B = -[W]^T \bar{I}_A \tag{3.26}$$

Note that from eqns. (3.5) and (3.6), the modified field patterns, $\tilde{e}$ and $\tilde{h}$, are frequency independent. The $[W]$ matrix is therefore also frequency independent. By changing the order of application of orthogonality relationships, it is possible to derive alternative equations in a similar way, but these can lead to an undetermined system of linear equations in some cases [Eleftheraides 1994].

g) Obtaining the GSM from the Mode Matching Equations

The GSM can be expressed in terms of the modal amplitudes as in eqn. (3.27).

$$\begin{bmatrix} \bar{b}_A \\ \bar{b}_B \end{bmatrix} = \begin{bmatrix} [S_{AA}] & [S_{AB}] \\ [S_{BA}] & [S_{BB}] \end{bmatrix} \begin{bmatrix} \bar{a}_A \\ \bar{a}_B \end{bmatrix} \tag{3.27}$$

The GSM matrices can be found by expressing the mode matching equations in terms of the modal amplitudes and applying the formulas from literature [Omar 1985]. The GSM matrices are expressed in terms of the $[W]$ matrix and the modal impedances and admittances as follows in eqns. (3.28) and (3.29).

$$\begin{aligned}[R] &= diag\left(\sqrt{\bar{Y}_A}\right)\cdot[W]\cdot diag\left(\sqrt{\bar{Z}_B}\right)\\ [T] &= diag\left(\sqrt{\bar{Z}_B}\right)\cdot[W]^T\cdot diag\left(\sqrt{\bar{Y}_A}\right)\end{aligned} \tag{3.28}$$

$$\begin{aligned}[S_{AA}] &= \left([R][T]+[I]\right)^{-1}\left([R][T]-[I]\right) & [S_{AB}] &= 2\left([R][T]+[I]\right)^{-1}\cdot[R]\\ [S_{BA}] &= [T]\left([I]-[S_{AA}]\right) & [S_{BB}] &= [I]-[T][S_{AB}]\end{aligned} \tag{3.29}$$

3.2.2 Convergence Tests and Verification

The formulation presented in the previous section has been implemented in MATLAB and number of tests were done to verify the program.

A check for the conservation of complex power is presented by Eleftheraides at al. [Eleftheraides 1994]. It checks that the power input is equal to the power output by verifying the relationship for propagating modes expressed in eqn. (3.30).

$$[S]\left[S^*\right] = [I] \tag{3.30}$$

Here, $[S]$ is the complete GSM for only propagating modes. This test was performed for various step geometries and at various frequencies and the program was found to be in agreement to within the limits of numerical accuracy.

Convergence tests gradually increase the number of modes retained and observe whether the entries of the GSM converge. They are useful both in verifying the formulation and in establishing an idea of the appropriate number of modes to retain for a specified accuracy. The results of a number of such tests for various geometries are summarized in figure 3.2. Note that GSM entries for single plane steps (b) and (d) appear to converge more quickly, whilst widely varying geometries (c) converge the slowest.

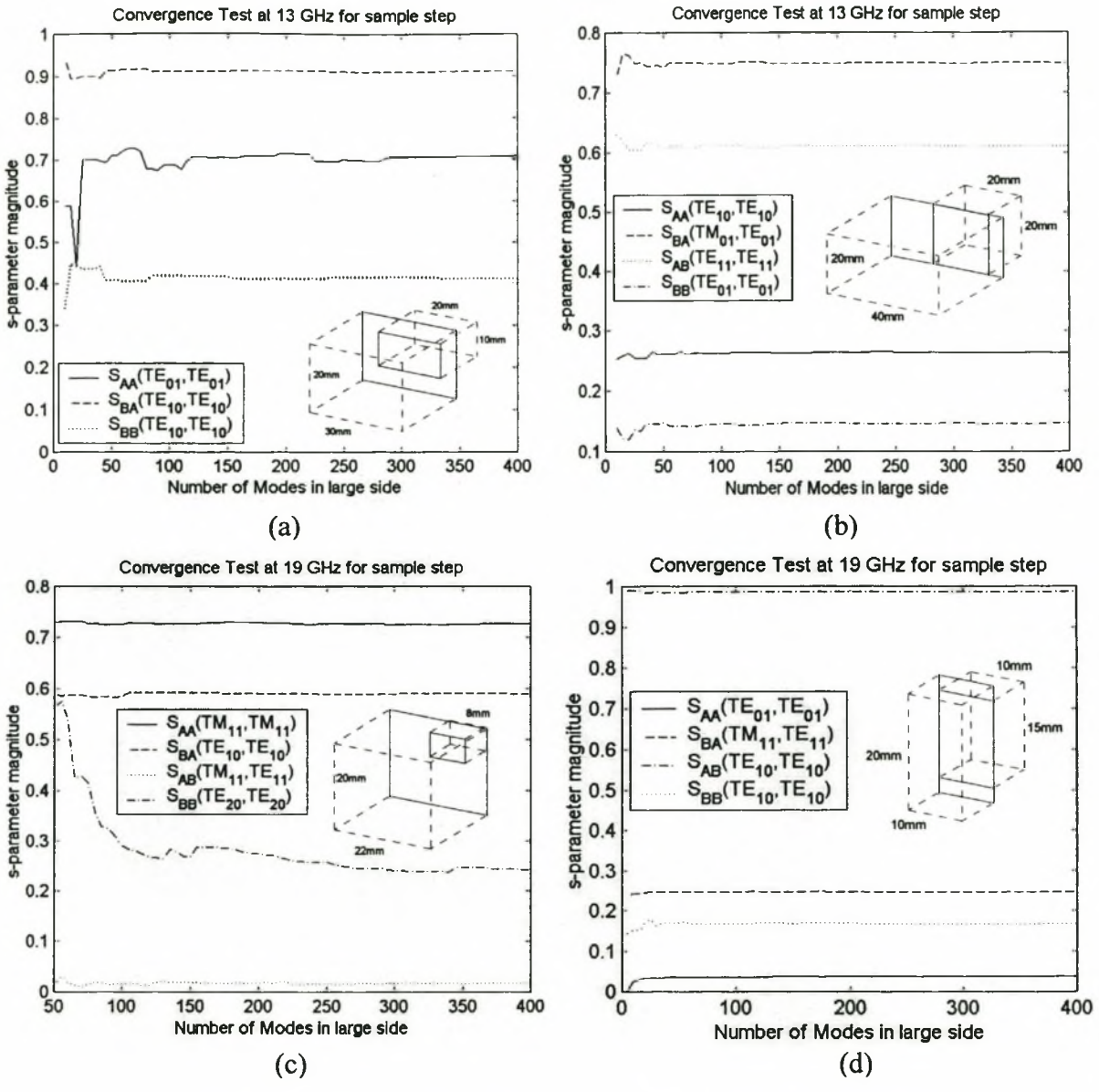


Figure 3.2 Convergence Tests for the Mode Matching Formulation

In addition to these tests, the program was tested against a commercial FEM code (Ansoft HFSS) and against another in-house mode matching code, which was written in the TE_x modal set formulation. In both cases, this formulation showed good agreement to the limit where the codes' results could be compared. The results of the comparison with the mode matching code are given in more detail in section 3.7

3.3 Cascading GSMs

In order to analyse systems of cascaded steps separated by uniform waveguide sections, the GSMs obtained by mode matching must be cascaded. This is done by progressively cascading adjacent GSM blocks until only one GSM block remains. The result is the GSM of the whole system. The derivation of the cascading formulas is readily available in the literature [Omar 1985] and only the formulas will be repeated here.

$$\begin{aligned}
[S_{AA}] &= [S^1_{AA}] + [S^1_{AB}][[I] - [S^2_{AA}][S^1_{BB}]]^{-1}[S^2_{AA}][S^1_{BA}] \\
[S_{AB}] &= [S^1_{AB}][[I] - [S^2_{AA}][S^1_{BB}]]^{-1}[S^2_{AB}] \\
[S_{BA}] &= [S^2_{BA}][[I] - [S^1_{BB}][S^2_{AA}]]^{-1}[S^1_{BA}] \\
[S_{BB}] &= [S^2_{BB}] + [S^2_{BA}][[I] - [S^1_{BB}][S^2_{AA}]]^{-1}[S^1_{BB}][S^2_{AB}]
\end{aligned} \tag{3.31}$$

In cases where a large number of modes are considered in the mode matching operation, and large GSM's must be cascaded, the direct application of the formulas above can become computationally expensive. Alessandri et al. [Alessandri 1992] propose that, since GSM blocks are typically separated by finite line lengths, and that over such distances, many non-propagating modes might be highly attenuated, it is reasonable to neglect those modes in the cascading operation. Such modes are termed 'localized' as their active influence is limited to a particular step or smaller part of the whole. When applied, this principle simply truncates the GSMs appropriately after the mode matching operation, and cascades the 'non-localized' modes as usual.

Another approach to easing the burden of cascading suggests the use of T-parameters, which, if applied to the formulation derived in section 3.2, would only require one frequency independent matrix inversion per step without having to perform the steps detailed in section 3.3. The main drawback of the T-parameter approach is its inherent numerical instability when dealing with non-propagating modes or with very high attenuations over physically long structures. This limitation can be overcome by a formulation that makes use of 'helping modes' (see Appendix 4), but this approach imposes a limit on the dynamic range of the analysis. An additional drawback is that T-parameters require equal numbers of electrical ports on either side of the discontinuity.

Alessandri et al. [Alessandri 1988] suggest a Y-matrix formulation that can be applied to good effect when cascading small to big, followed by big to small steps, but equivalent Z-, G- and H- parameter formulations must be used in other step configurations to create the same general effect. The effort of combining the different formulations for a general cascade exceeds the gains made by their independent contributions.

To conclude, the most reliable and generally applicable cascading approach appears to remain the GSM formulation. The concept of localized modes can be used to good effect to reduce the cascading effort in the context of this approach.

3.4 An Equivalent Circuit Formulation for Mode Matching

Some references to equivalent circuit representations of the mode matching formulation exist in the literature [Weisshaar 1996, Hsu 2000]. Motivations for this effort are to gain insight into the physical nature of mode matching and to identify ways of approximating the formulation with circuit theory. In addition, such representations can provide important information regarding the electrical behaviour of waveguide discontinuities, such as pursued in the classic paper by Eisenhart and Khan for the circuit impedance of shunt mounted devices in waveguide [Eisenhart 1971]. Perhaps the most straightforward mapping from the mode matching equations to circuit theory is by Weisshaar et al., where voltage controlled voltage sources and current controlled current sources directly implement the equations derived in section 3.2. Later work by Hsu et al. applies only to H-Plane steps and introduces the concept of a 'vector' multiport transformer that couples the modes on either side of the discontinuity. A more complete equivalent circuit formulation is introduced here [Vale 2001]. This circuit applies to all mode matching problems that can be expressed in terms of the mode matching equations from section 3.2. The circuit uses only ideal 2-port transformers and transmission lines, but exactly solves the equivalent mode matching problem in circuit theory. Its development follows from the mode matching derivation in section 3.2.

Consider a single mode on the A side, and a number of modes $1..M$ on the B side of a discontinuity. The mode matching equations can be rewritten as in eqns. (3.32) and (3.33).

$$V_{A1} = W(1,1)V_{B1} + W(1,2)V_{B2} + \ldots + W(1,M)V_{BM} \tag{3.32}$$

$$\begin{aligned} I_{B1} &= -W(1,1)I_{A1} \\ I_{B2} &= -W(1,2)I_{A1} \\ &\ldots \\ I_{BM} &= -W(1,M)I_{A1} \end{aligned} \tag{3.33}$$

If $M = 1$, then a 2-port transformer, as shown in figure 3.3, could model the equations exactly if the winding ratio, n, is chosen as $W(1,1)$. In the case where $M > 1$, however, the A-side voltage can be seen as a sum of transformed voltages – or a series connected circuit. This observation prompts the circuit segment in figure 3.4 as a potential part of the full equivalent circuit. This circuit also satisfies the desired current relationships in eqns. (3.33).

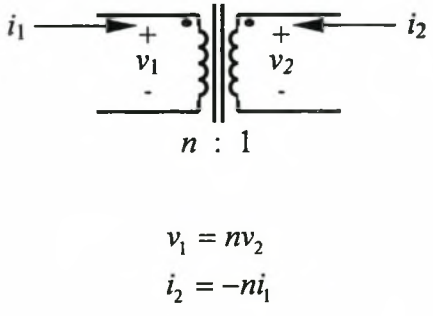


Figure 3.3 An ideal transformer

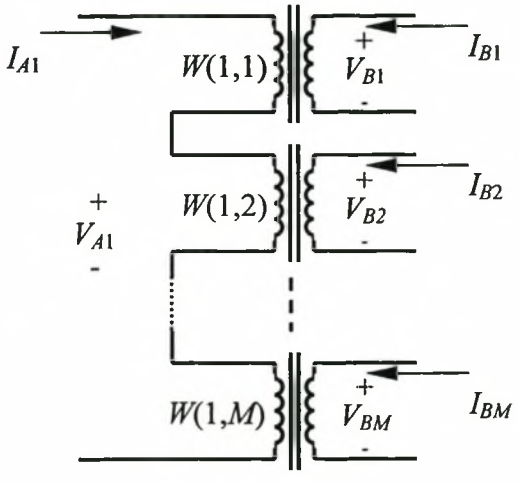


Figure 3.4 Circuit fragment for N=1, M>1

If, in a similar way, a single mode on the B-side is matched against a number of modes on the A-side, as before, the mode matching equations are rewritten in eqns. (3.34) and (3.35)

$$\begin{aligned} V_{A1} &= W(1,1)V_{B1} \\ V_{A2} &= W(2,1)V_{B1} \\ &\ldots \\ V_{AN} &= W(N,1)V_{B1} \end{aligned} \tag{3.34}$$

$$I_{B1} = -W(1,1)I_{A1} - W(2,1)I_{A2} - \ldots - W(N,1)I_{AN} \tag{3.35}$$

The summation of currents in eqn. (3.35) indicates a parallel connected circuit fragment, shown in figure 3.5. As before, this circuit also satisfies the voltage relationships in eqn. (3.34).

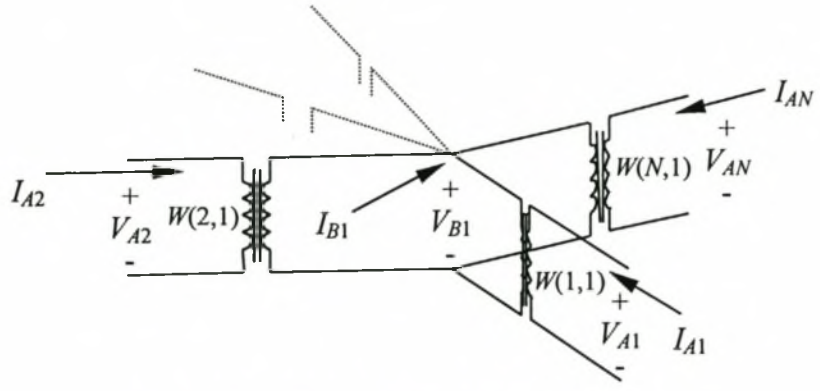


Figure 3.5 Circuit fragment for $N>1$, $M=1$

The circuit fragments in figures 3.4 and 3.5 are special cases of the sought-after general equivalent circuit that would hold for $N \geq 1$ and $M \geq 1$, and can be combined in a general circuit as shown in figure 3.6.

The generalized scattering matrix formulation assumes all ports to be terminated in infinite lines, or by the modal wave impedance. The latter representation is used for the circuit in figure 3.6 as it allows a more convenient way of visualizing the circuit. In cascade with other mode matching discontinuities, the ports would not be terminated with lumped elements, but would instead extend as transmission lines over the length of waveguide separating the discontinuities. The propagation constant of these lines would correspond to the modal propagation constants as given in eqns. (3.3).

It is interesting to note the similarity between the circuit in figure 3.6 and the equivalent circuit of a post mount in waveguide, as determined by Eisenhart et al. Both consist of a compound structure of simple two-port transformers that can grow indefinitely as higher order effects are taken into account. The increased complexity of the post mount problem, however, requires a more intricate determination of the lumped impedances attached to the transformer terminals in Eisenhart's model and this requires the use of approximations.

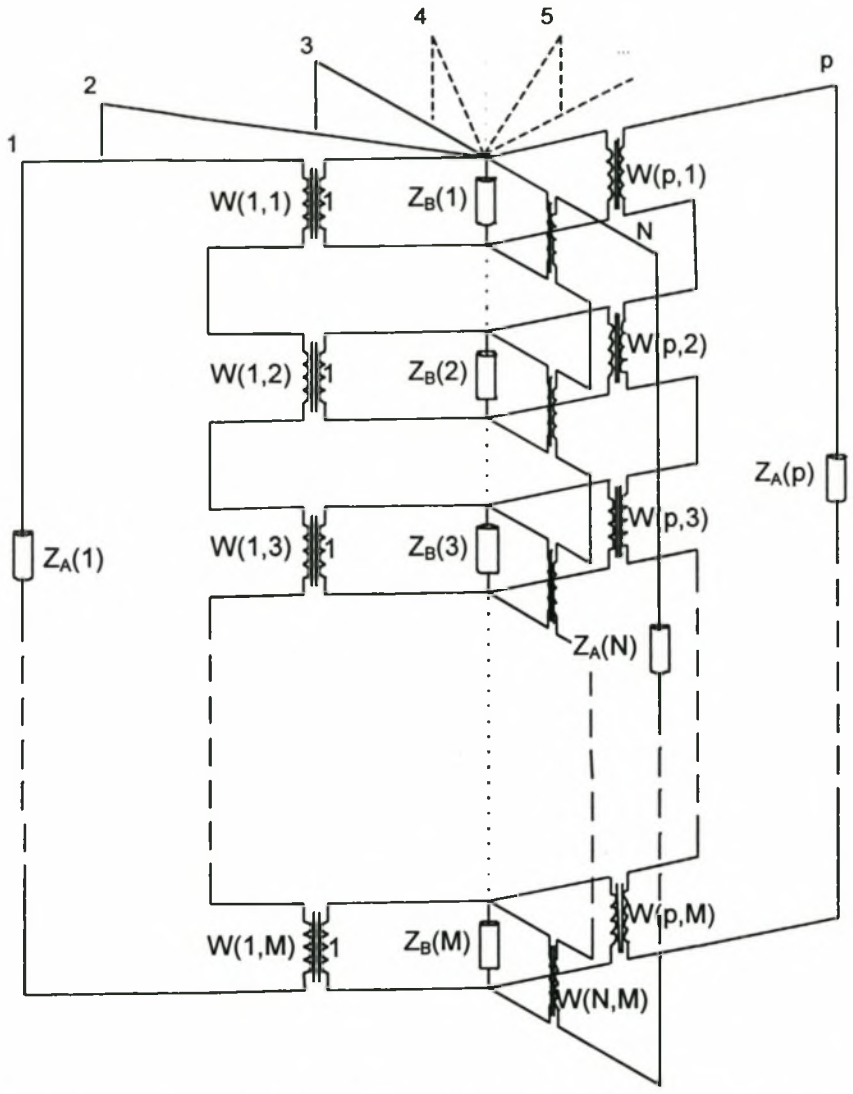


Figure 3.6 Lumped element - terminated Equivalent Mode Matching Circuit

In order to verify that the equivalent circuit models exactly the mode matching formulation, it is necessary to prove that s-parameter relations for the circuit (calculated using circuit theory) correspond to those derived in section 3.1. This is easily done by examining the scattering parameters between any two un-terminated ports of the equivalent circuit. If these two ports are extended as transmission lines, whilst the others are terminated with lumped impedances, then it can be shown that the corresponding entries of the mode matching GSM can be derived from circuit theory applied to the equivalent circuit. If the entire GSM of the circuit is solved in an equivalent manner, the formulation reduces to the exact equations for the GSM of the mode matching problem, verifying that solving for the GSM of the circuit can be seen as an equivalent mode matching operation.

Interesting parallels can be found between approximations that have been widely used in practice and similar approximations that can be made of the equivalent circuit. In dominant TE_{10} propagation, it is a common to assign an equivalent circuit impedance proportional to the inverse of the height of the waveguide. In the case of E-plane steps, the resulting reflection that occurs at the step can be approximated as a result of an impedance variation over the step. According to the equivalent circuit formulation above, however, there is no variation in impedance for the TE_{10} mode across the E-plane step due to the choice of voltage and current definitions. Consider now a TE_{10} to TE_{10} mode matching operation done on an E-plane step. With the equivalent circuit formulation this reduces to two equal impedance lines coupled by a transformer with turns ratio equal to $W(1,1)$. Scattering that occurs at the step is due to a transformed apparent impedance. In figure 3.7, the value of $W(1,1)^2$ for various geometries is plotted against the E-plane height ratio for a centred step. The agreement between the values indicates that the standard height ratio approximation is equivalent to a mode matching operation involving only the dominant mode.

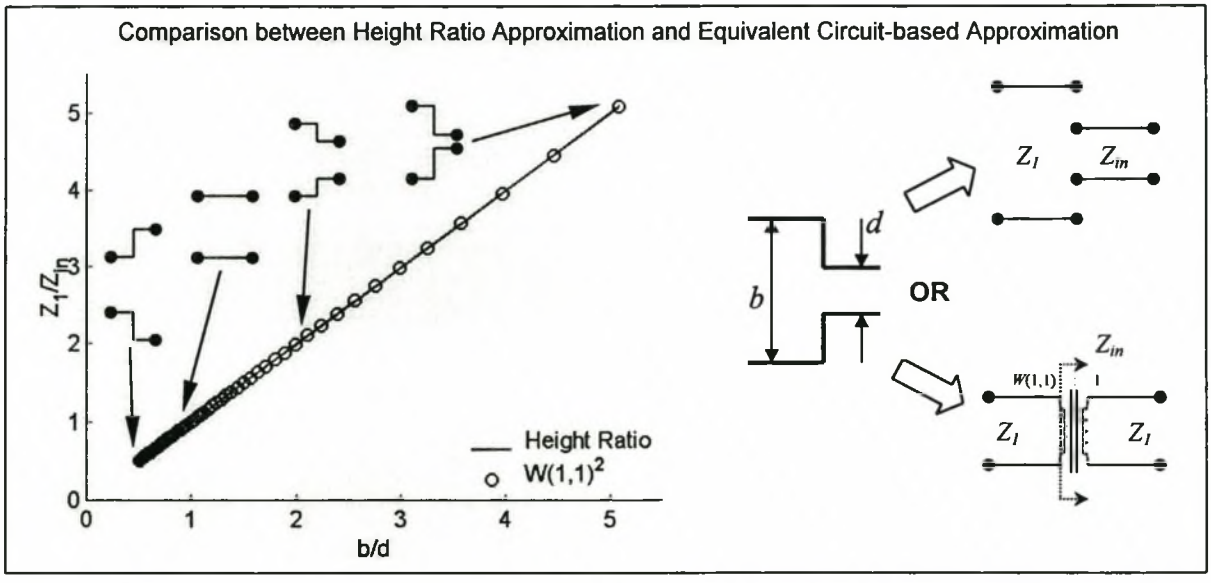


Figure 3.7 The equivalence of the standard TE_{10} height-impedance approximation and the new equivalent circuit single-mode approximation.

3.5 Automated Intelligent Mode Selection for Fast Mode Matching

The computational effort in mode matching arises from the matrix operations that have to be performed to obtain the GSM of the discontinuities as seen in eqns. (3.28) and (3.29). In order to reduce the effort of these computations, whilst retaining the accuracy of the analysis, standard practice is to use knowledge of modal behaviour to predict which modes might not interact with a particular geometry and to neglect them in the mode matching operation. A typical example is the use of only TE_x modes for the analysis of TE_{10} excited steps.

While this type of approach does result in more reasonably sized problems, it has certain drawbacks. Firstly, it cannot be used with a great deal of consistency in cascaded systems of steps that do not all adhere to the symmetry conditions. Secondly, the calculated results of such an analysis are made in terms of the alternate mode set, requiring conversion to the more standard TE_z/TM_z formulation [Vale 2000b]. Thirdly, the great flexibility of the mode matching method is compromised by the limitations on step geometry and modal excitation imposed by the approach. Finally, it does not lend itself well to generalization; each problem must be separately considered by the user and an applicable code written for it. Other strategies rely on similar empirical techniques and suffer from the same drawbacks. They typically require the user to use symmetry conditions to manually exclude certain modes from the operation, provided he/she is confident that they do not interact with the problem at hand. While this can be done, it is difficult to implemented to its fullest possible extent and often prohibitively difficult, such as with cylindrical to rectangular waveguide steps.

This section will exploit the equivalent circuit introduced in the previous section to identify a new strategy for mode selection in mode matching problems [Vale 2001]. This strategy greatly reduces the number of modes needed to do an accurate analysis of certain geometries and can be seen as an alternative to the strategy of truncating the modes at a particular value of k_c as described in section 3.2.1. The technique can be used to systematically reduce the mode matching content to contain only the critical interacting modes for a particular problem. It can also be used to solve problems of cascaded steps with significant improvements in speed over the conventional approach. Finally, excellent physical insight can be gained in an easy manner. The approach does not rely on the use of alternate mode sets and does not limit

the modal excitation. It can also be used for (and has been tested with) a variety of other types of discontinuities, such as rectangular to circular waveguide steps. In the first part of this section, the technique will be described for a step discontinuity in rectangular waveguide and compared with the standard technique. The second part will show how the technique can be used to solve large cascaded systems.

3.5.1 Subcircuit Reduction

Consider a transformer from the equivalent circuit in figure 3.6. In this case the transformer, shown in figure 3.8a, has a turns ratio, $W(m,n)$. If this value becomes very low, the low voltage side of the ideal transformer occurs in series on the A-side and the low current side occurs in shunt on the B-side. As $W(m,n)$ is reduced to zero, for moderate terminating impedances on either side, the transformer will look increasingly like a short circuit from the A side and an open circuit from the B side. In the limit where $W(m,n)$ tends to zero, the transformer would appear as shown in figure 3.8b.

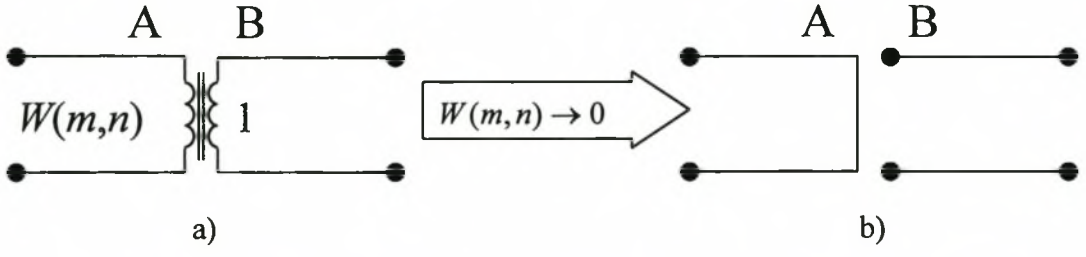


Figure 3.8 The effect of a low value of $W(m,n)$ on the coupling transformer

Now, for ease of illustration, consider a very small mode matching example with 2 modes on side A and 3 modes on side B, as shown in figure 3.9a. If a sample $[W]$ matrix with certain entries equal to (or very close to) zero applies to the circuit then, according to the observation above, the circuit could be redrawn as shown in figure 3.9b. This circuit can then be pulled apart to yield the two separate, independent subcircuits in figure 3.9c. Since the $[W]$ matrix is frequency independent, this holds for any frequency that might be considered.

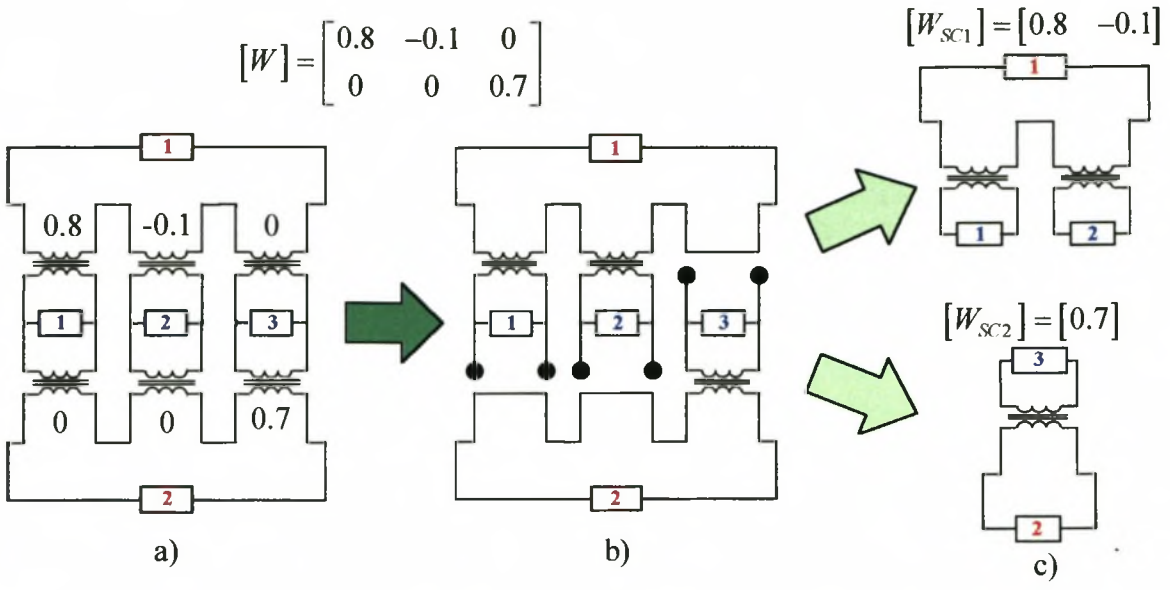


Figure 3.9 Reduction of a MM problem into independent sub-problems

The great advantage of this approach is that it automatically identifies which modes are relevant to a specified excitation. If, for example, only TE_{10} mode excitation is of interest, only those modes that actually interact (couple) with the TE_{10} subcircuit need be considered.

In some systems, where the $[W]$ matrix is particularly sparse, the reduction described above can lead to a large number of independent circuits, each with relatively few modes involved. If solved independently, the resulting GSM entries will match the relevant corresponding entries of the complete GSM that would result if the standard mode matching procedure were applied to the problem provided only transformers were 'neglected' that had a zero magnitude turns ratios. Typically, sparse $[W]$ matrices arise when step geometries are simple or symmetrical. Such situations (which are common) can yield great savings in computational effort with the use of this selection strategy. Figure 3.10 shows the analysis of a simple step for only the TE_{10} mode, and compares it to standard full-mode analysis. The agreement between the selected parameters calculated by the two approaches is easily observed.

Asymmetric, complex geometries typically yield very few, if any, zero-entries in the $[W]$ matrix. In such cases, there is typically no breaking down of the mode matching problem, and the exact strategy for simple steps described above reduces to the standard practice of retaining all available modes. By approximating suitably small values of $W(m,n)$ as zero,

however, the strategy can be made to discard weakly coupled modes in favour of more strongly coupled modes. This is done by investigating the coupling ratio shown in eqn (3.36).

$$R_c(m,n) = \sqrt{\frac{Z_m^{(A)}\left(f_{\text{maxrat}}\right)\cdot W\left(m,n\right)^2}{Z_n^{(B)}\left(f_{\text{maxrat}}\right)}} \tag{3.36}$$

Here f_{maxrat} is that frequency in the analysis sweep that maximizes R_c. If this lies below a predefined threshold value, e.g. $R_c(m,n) < 0.01$, then the coupling is discarded and the transformer turns ratio is assigned zero. Independent subcircuits extracted with this approach will not yield reduced GSMs in full agreement with the full mode analysis, but, since the strategy discards modes based on coupling strength, rather than on the value of k_c, it is expected to produce more accurate results for the same number of modes than the standard full mode approach, which might truncate a higher order mode with significant coupling. Figure 3.11 illustrates this with the reduced mode analysis of an asymmetrical step. The much smaller reduced system is observed to be more accurate than the same sized full mode matching approach with modes ordered according to cut-off frequency. This inexact reduction strategy was successfully tested in a rectangular-to-cylindrical step environment, with observed reductions of the order of 75% of the full problem with R_c threshold at 0.001. Minimal degradation of analysis accuracy was observed.

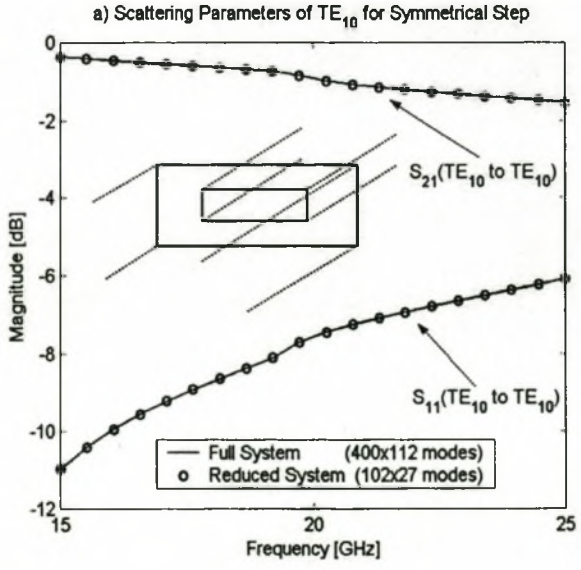


Figure 3.10 Exact reduction of a symmetric step for the TE_{10} mode.

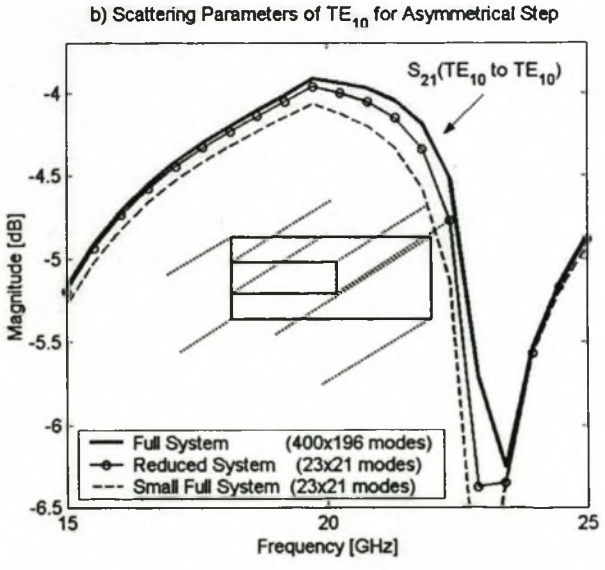


Figure 3.11 In-exact reduction of an asymmetric step for the TE_{10} mode.

3.5.2 Fast Cascade Structure Analysis with the aid of Reduced Mode Matching (RMM)

The advantages of the intelligent mode selection scheme's ability to divide mode matching problems into smaller independent subcircuits are best seen when analysing cascaded discontinuities. The following procedure exploits the reduction technique and the concept of 'localized modes' [Alessandri 1992], and provides an efficient way of cascading step structures in waveguide.

1) Each step comprising the structure has its full W-matrix calculated in the usual way.

2) Each equivalent circuit is divided into all its independent subcircuits as described above. The subcircuits are indicated by a column of dots aligned with their corresponding step in the diagram alongside

3) Subcircuits are now connected on the basis of mode-sharing. If adjacent subcircuits share the same mode, and that mode can sufficiently propagate over the length of waveguide separating the steps, then the subcircuits are considered 'connected'. The connections are indicated by blue and red lines in the graph.

4) Finally, based on the desired port excitation, the relevant subsystem is traced and then cascaded using the GSM formulation (with localized modes). In this example, TE_{10} response is desired. The red lines and dots indicate the TE_{10} subsystem. Its response is compared to the standard analysis alongside.

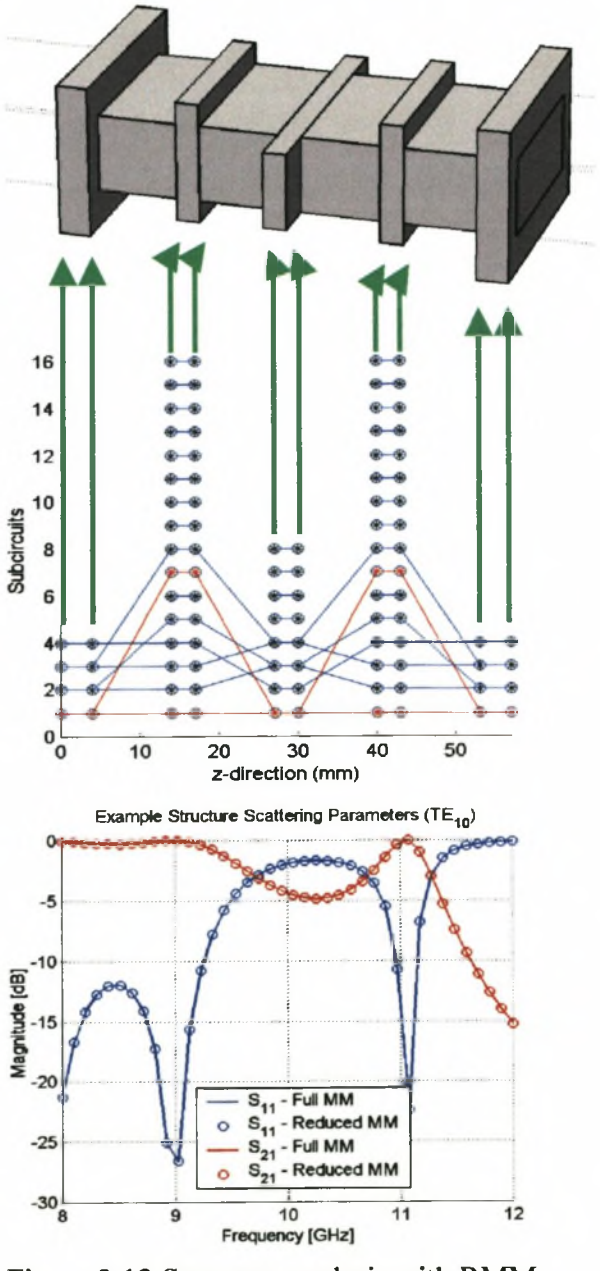


Figure 3.12 Structure analysis with RMM

3.5.3 Verification of the New Analysis Technique

The analysis strategy described above was tested with a number of sample waveguide structures and the results and execution times compared with a standard full mode matching/GSM analysis.

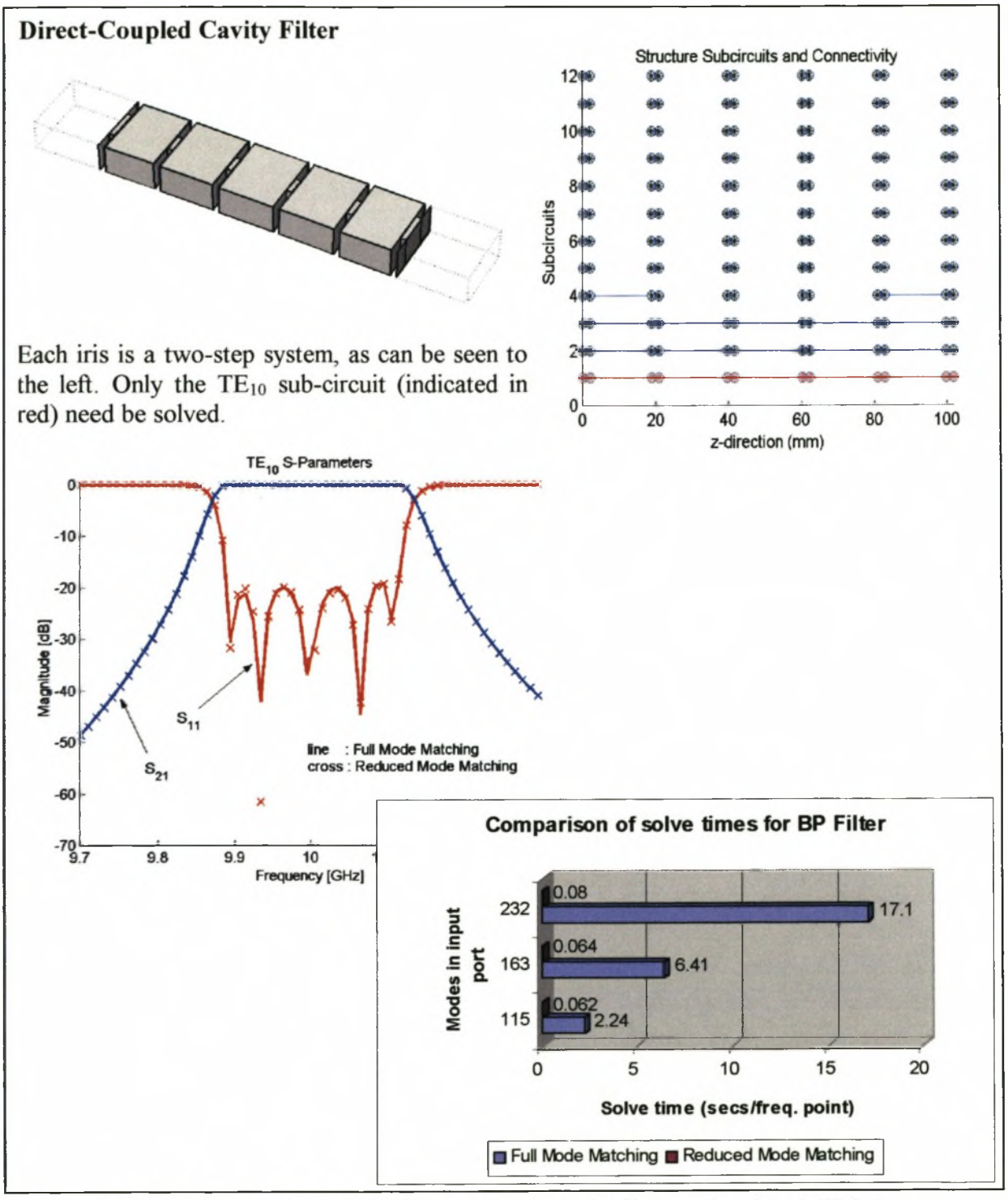
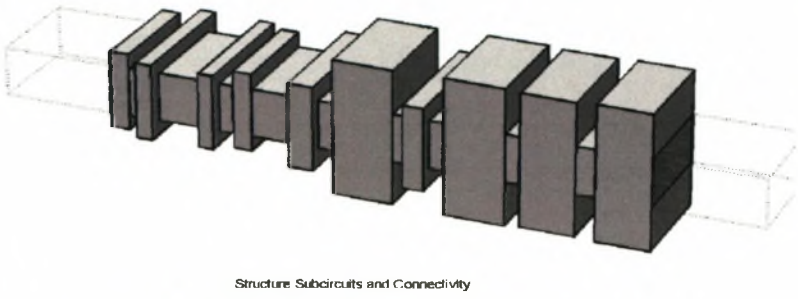


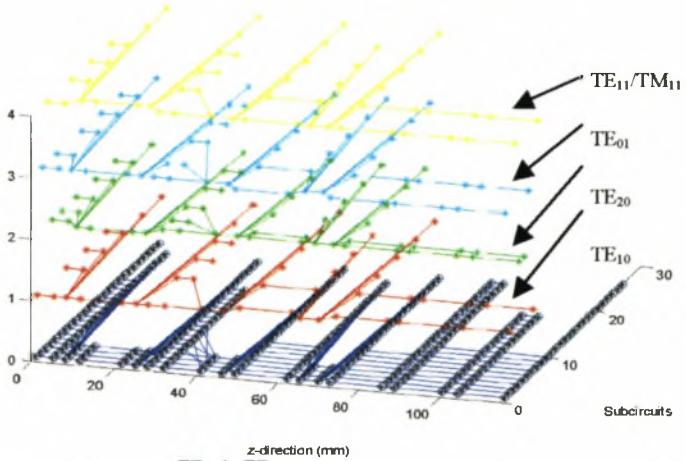
Figure 3.13. Example 1. A direct coupled cavity filter analysed with RMM

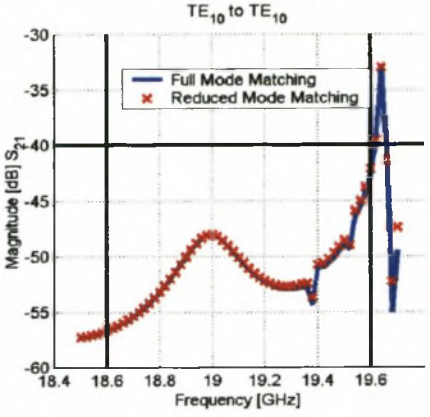
Multimode Bandstop Filter

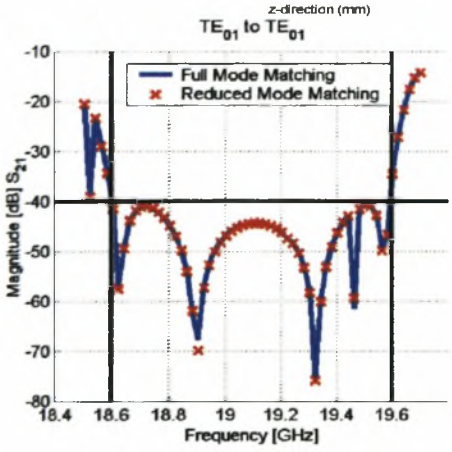


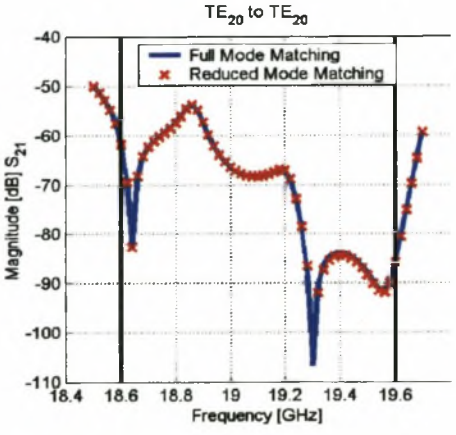
Structure Subcircuits and Connectivity

Scattering parameters for all 5 propagating modes are desired. The new technique resolves the structure into 4 independent subcircuits for these modes. These can each be separately solved for the various propagating modes.









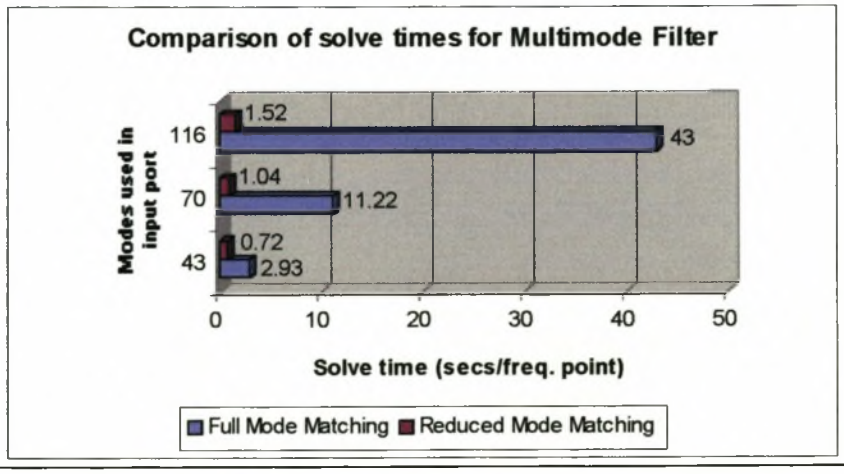
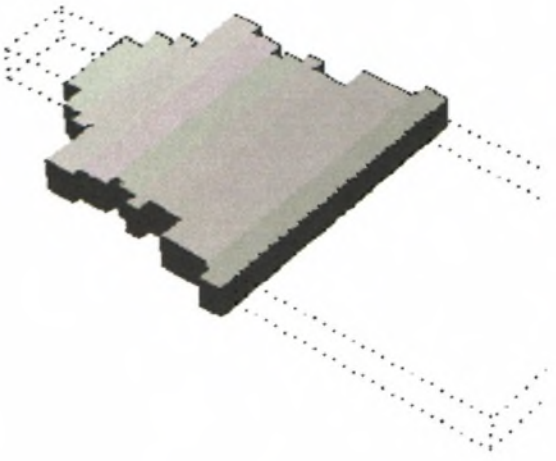
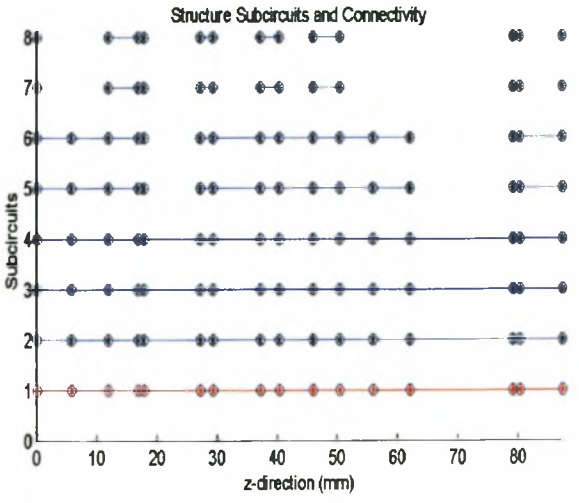


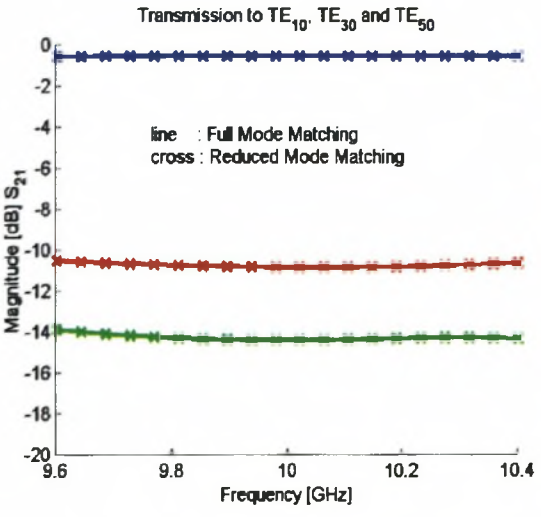
Figure 3.14. Example 2. A Multimode Bandstop Filter (see chapter 5) analysed with RMM

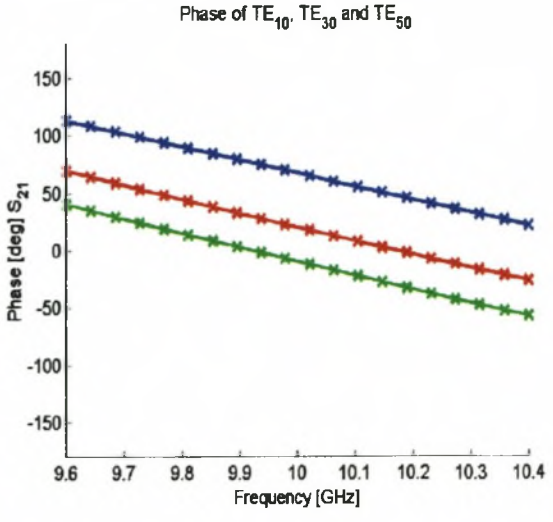
Multimode H-Plane horn





The horn is excited by the TE_{10} mode. The sum of TE_{10}, TE_{30} and TE_{50} modes on the large side form a desired field pattern. The solver extracts the TE_{10} sub-circuit (shown in red in the subcircuit map) and ignores the other sub-circuits.





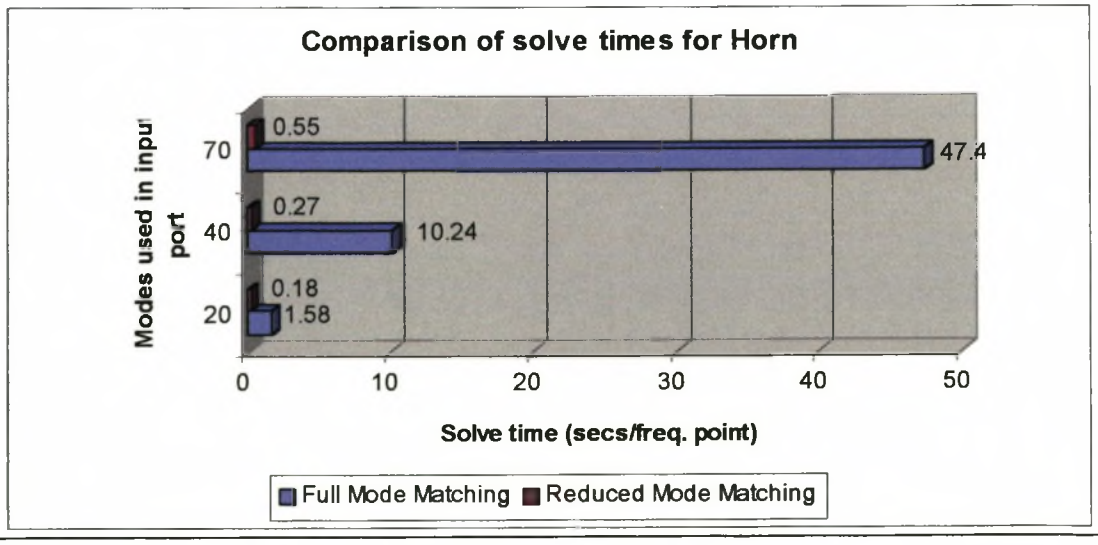


Figure 3.15. Example 3. A Multimode Horn (see chapter 6) analysed with RMM

All the examples above were analysed with the same analysis code, which uses exact reduction of subcircuits (R_c set to 10^{-10}). Localized modes are discarded after 40dB attenuation. The examples above show good agreement between the new reduced mode matching-based cascaded analysis and the traditional full mode matching/GSM analysis. Discrepancies that occur are caused by the discarding of localized modes and disappear if the rejection threshold is increased sufficiently (at the cost of slightly increased computational effort). The massive gains in analysis speed are apparent in each example.

The implementation of reduced mode matching described above requires the full computation of the *W*-matrix for all the modes *before* reduction. Since the *W*-matrix is frequency independent, this is typically a relatively cheap operation in comparison to the computational efforts of performing a frequency sweep. This is especially so in the case of rectangular waveguide, where the W-matrix entries are trivial to compute. In cases where numerical integration is used to determine the W-matrix entries, the computation of an entire matrix can become overwhelming. In such cases, a basic modification to the RMM formulation can be applied that builds up the essential elements of the W-matrix as subcircuit extraction progresses, resulting in considerably fewer evaluations of W-matrix entries. The modified approach can be applied to both single step analysis and cascaded discontinuity analysis just as the formulation described here. Since this investigation concerns itself only with rectangular waveguide discontinuities, the formulation developed here is, however, adequate. The development and testing of an algorithm that accomplished the modified formulation will not be presented here and is left as a subject of further research.

This section has introduced a new strategy for rapid analysis of cascaded step structures in waveguide. This allows much faster analysis of a great variety of arbitrary structures with arbitrary excitations than was originally possible with the standard MM/GSM technique. For the purposes of CAD, such improved analysis techniques enable more sophisticated and more generally applicable design techniques to be developed. In addition to these algorithms, the speed of analysis can be advanced even more by the use of so-called surrogate models.

3.6 Surrogate Models for Mode Matching

The concept of using a surrogate model to substitute a full EM analysis model has been extensively covered in the literature [Bakr 2000, Watson 1997, Gupta 1998]. In essence, it proposes a general purpose model that can be 'trained' by the full EM model at a number of points (dimensions or frequency), that would ideally be far fewer than would be expected to be used by the design or optimization process. Instead of referring to the full EM model, the optimisation or design process would call the surrogate model, which would be expected to give a reasonable estimate of the true response. Both interpolators and neural networks have been investigated as potential surrogates with varying success. A very promising example of these investigations [Lehmensiek 2001], describes a multivariate, multiple output adaptive interpolation scheme that requires remarkably few points to establish.

It is, however, not the purpose of this dissertation to further meaningfully the field of surrogate models. Accordingly only a very simple, multivariate, multiple output, interpolation scheme is described here.

3.6.1 Multiple Output Oriented Interpolation scheme

This scheme is based on the simplex-coordinate system, widely used in FEM, to provide a robust and very fast interpolation technique for multiple output functions. Such multiple output modelling is important in multimode applications because there are typically a large degree of electrical parameters necessary to describe the discontinuities. The limitation of this scheme is that it requires a grid of data points in n-dimensions (for n input variables), though the grid does not have to be uniform. The two-variable case will be used to describe its operation.

The scheme is based on the proposal that the interpolated value, f^I, at a point surrounded by a N other sample points, f^R, can be described as in eqn. (3.37).

$$f^I\left(x_1, x_2, x_3, \ldots\right) = \sum_{i=1}^{N} w\left(x_1, x_2, x_3, \ldots; x_1^i, x_2^i, x_3^i, \ldots\right) \cdot f^R\left(x_1^i, x_2^i, x_3^i, \ldots\right) \tag{3.37}$$

Since the weighting function, w, is not a function of the actual functional values, f^R, and only a function of the position of the point, the same weighting function can be used for any number of arbitrary output functions. This makes the scheme very easy and quick to apply to multiple-output interpolation, as the weights are simply re-used for each output. The nature of the interpolant will depend on the number of points, N, considered and the nature of the weighting function, w.

This scheme proposes that in a rectangular grid of points, only those points forming a hypercube about the interpolated point are to be considered, and that the weighting function should weight those points in relation to the proximity of the interpolated point. For the two-variable case, the weighting scheme is shown in figure 3.16. The weight of a particular point is assigned as a ratio of the opposite sub-area to the total area as indicated in the figure.

This scheme provides a smooth transition between the surrounding points, with exact agreement at the sample points, as well as reduction to a linear interpolation at extremities of the hypercube. Figure 3.17 shows interpolation of one cell for two-variables where this behaviour is easyly observed. Figure 3.19 shows interpolation of a two-variable function for two grid densities and the associated relative error for each case. The original function is in figure 3.18.

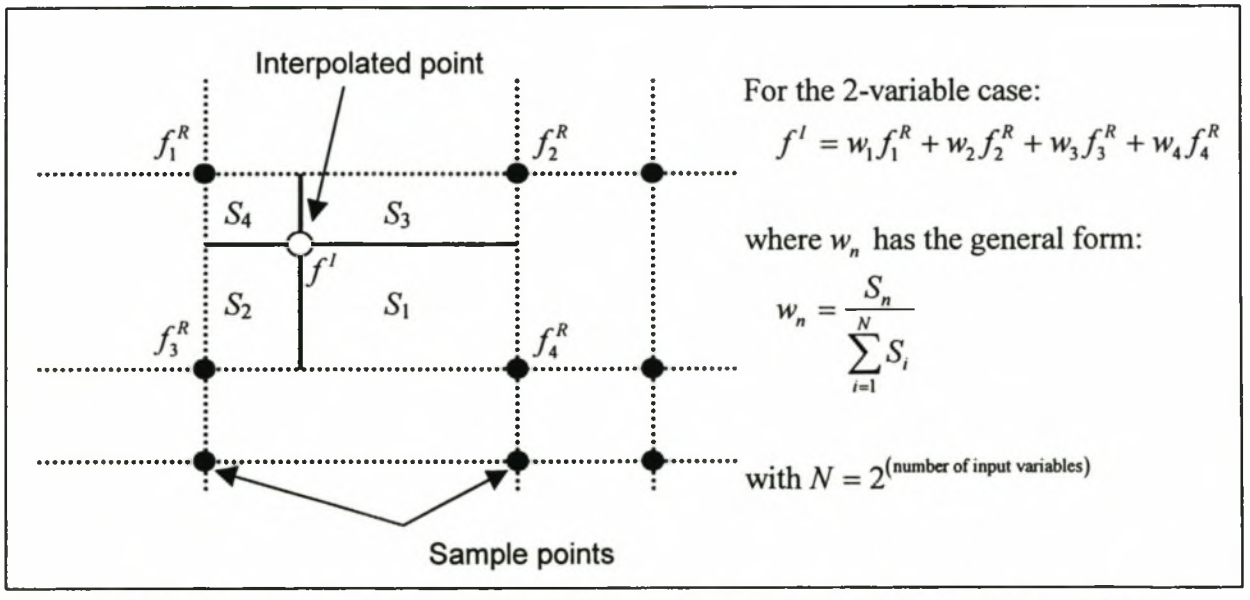


Figure 3.16 The weighting scheme of the interpolation technique for two variables

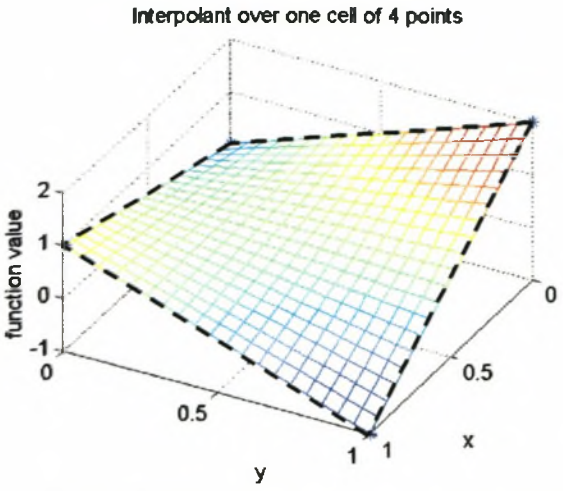


Figure 3.17 Interpolant over a single cell

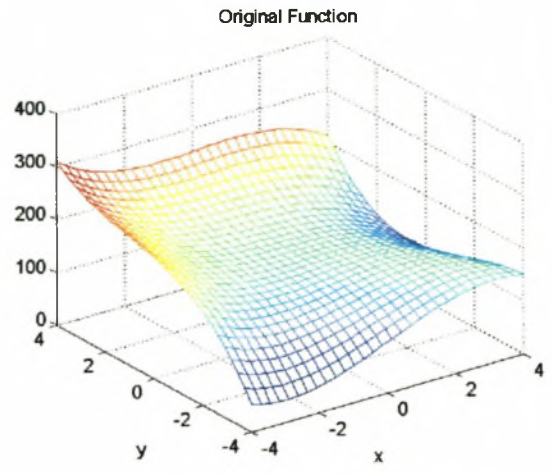
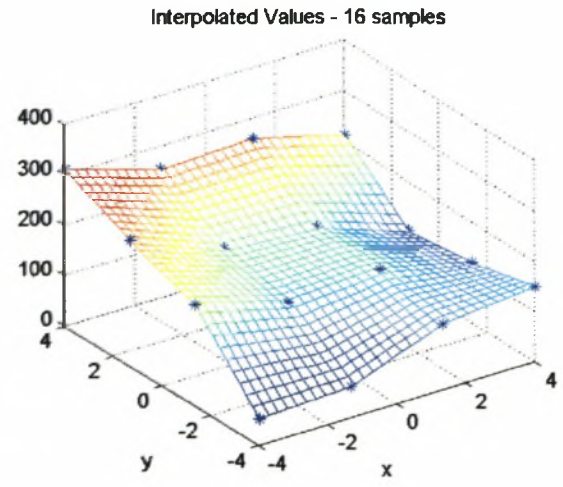
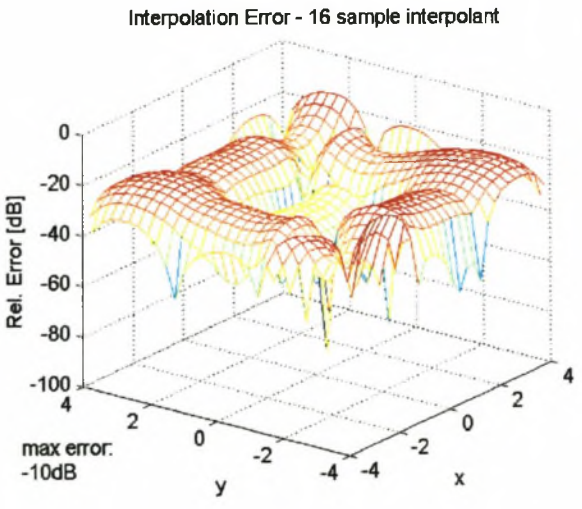
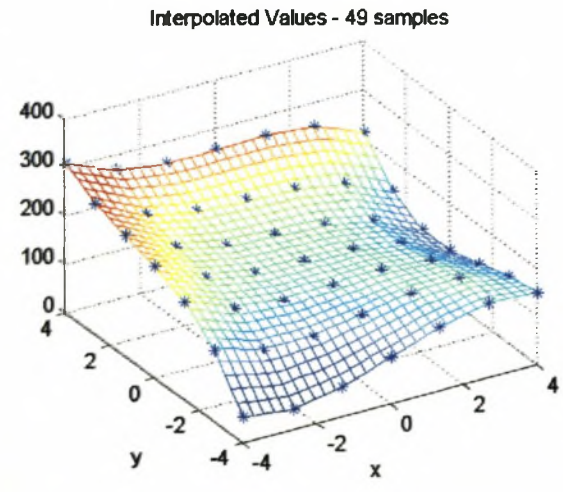


Figure 3.18 Two-variable function used to evaluate interpolation scheme







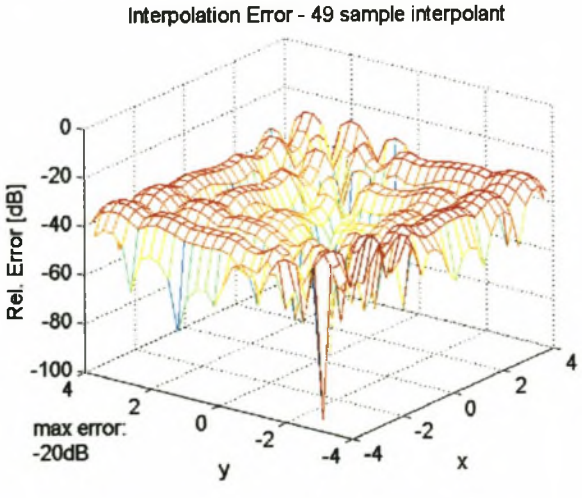


Figure 3.19 Interpolated function values and relative error for two grid densities

3.6.2 Surrogate Model Application to EM analysis

In order to illustrate the potential for CAD use of the surrogate model scheme, the direct-coupled cavity filter example (redrawn in figure 3.20) of section 3.5.3 is revisited. In this analysis, an interpolation-based model is used to replace the EM model for an H-plane iris. The variables are the aperture width and aperture length. First a uniform $n \times n$ grid of sample aperture widths and lengths is defined with grid extremes as shown in figure 3.22.

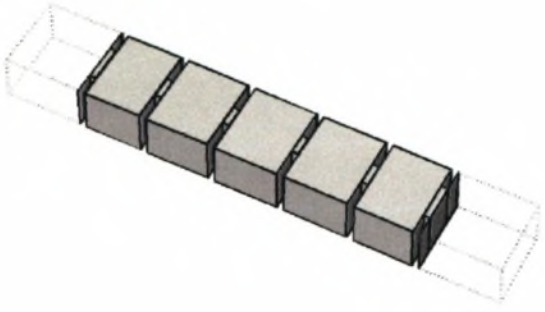


Figure 3.20 Structure of the filter

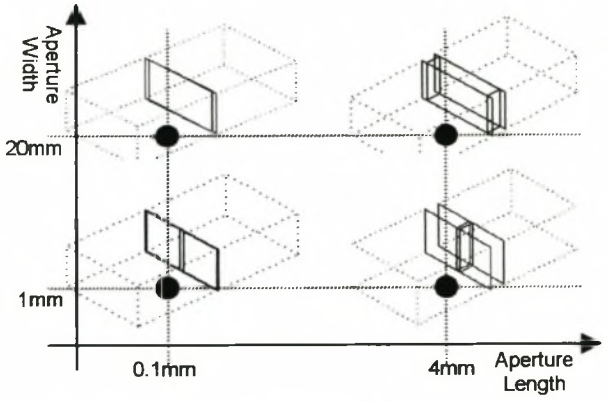


Figure 3.21 Extremes of the sample grid

Each sample point in the grid is analysed with a full frequency sweep using the fast analysis method. The resulting GSMs are truncated to contain only TE_{10} mode information and stored as data points for the interpolator. Analysis of the filter is then accomplished by cascading TE_{10} GSM data from the interpolation based model for each iris. Figures 3.22 and 3.23 show scattering parameter results of the analysis with a 7×7 grid (49 samples) which took 45 seconds to generate and a 10×10 grid (100samples).

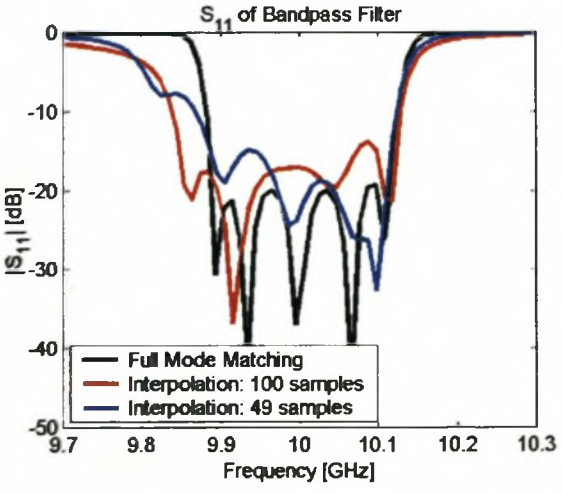


Figure 3.22. S_{11} (TE_{10}) of the filter

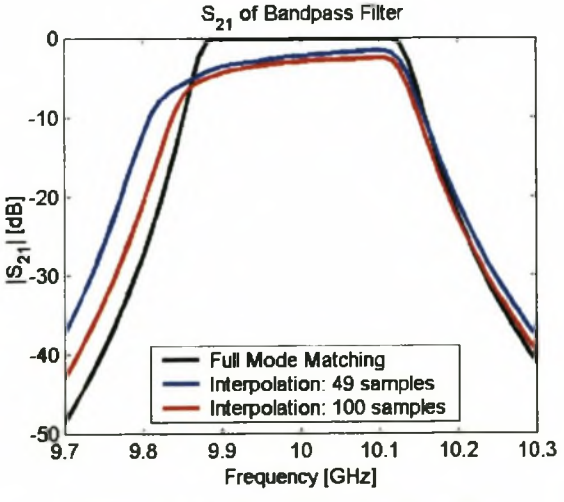


Figure 3.23. S_{21} (TE_{10}) of the filter

Clearly, increasing the grid density, either by the addition of more sample points, or by a more reasonable choices of grid extremes, will lead to a more accurate analysis of the structure. Note, however, that error with the interpolation-assisted analysis arises in two ways. First, from interpolation error and then from GSM truncation error, where potentially contributing modes other than TE_{10} are not taken into account. The costs of including a larger GSM are increased memory requirements, as well as increased computational effort in cascading the GSMs. The cost of reduced accuracy, however, is offset by the significant gain in analysis speed, shown in figure 3.24. The interpolation assisted model provides almost an 8-fold increase in analysis speed over even the fast mode matching analysis from section 3.5. Note that this graph is drawn with a logarithmic scale.

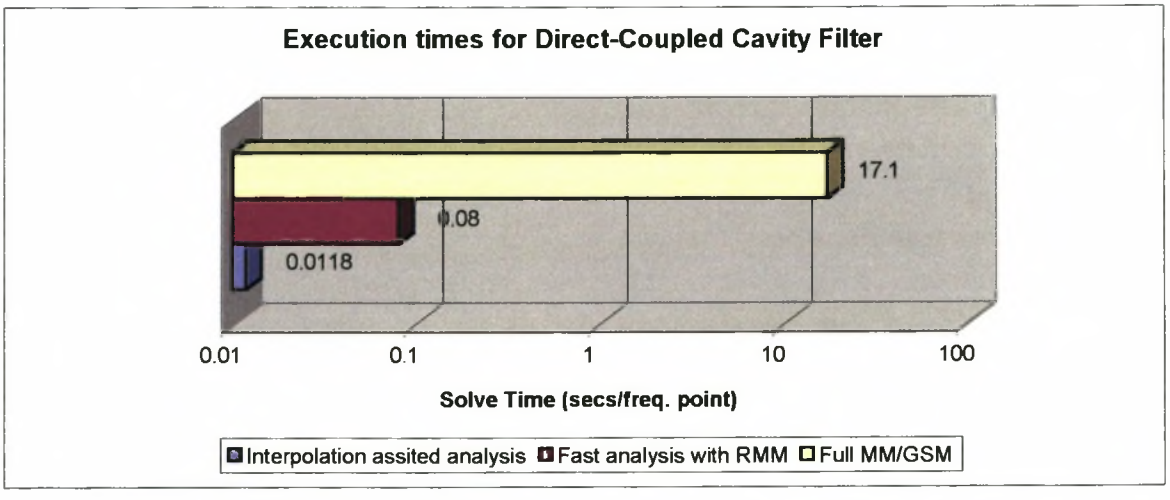


Figure 3.24 The benefit of the interpolation based model in terms of speed

A basic interpolation scheme has been devised that can be exploited as a surrogate model for mode matching. It provides very quick multiple-output interpolation and is robust. Its operation in the proposed CAD scheme would be as a alternative to a full EM model of various functional blocks, described in chapter 2, or as the coarse model in a space mapping optimization or design procedure [Bandler 1995]. The potential advantage of using these kinds of approximate functional blocks is much faster analysis at the cost of accuracy and a fixed time investment for training the surrogate model.

3.7 Modal Transforms in Rectangular Waveguide

As discussed in section 3.2, solutions to the Helmholz equations in rectangular waveguide can be expressed in terms of the well-known transverse electric (TE_{mn}) and transverse magnetic (TM_{mn}) modes, with the total field distribution expressed as an infinite sum of these modes. The direction to which the E or H fields are transverse, can be any of the *x*-, *y*- or *z*-directions, leading to three variant solutions, namely the TE_z/TM_z , TE_y/TM_y and TE_x/TM_x mode sets. While the TE_z/TM_z formulation is the most frequently used, the other modal sets can be of use in certain situations, such as in the development of specialized mode matching formulations as discussed earlier. Their application to this dissertation, however, lies mainly in the development of multimode measurement techniques to be presented in the next chapter. The ability to convert GSMs between different modal set representations is essential to the development and implementation of these techniques.

The derivation presented here yields simple matrices for the transformation between the generalized scattering parameters of the three mode sets. The use of the matrices is similar to those used for geometric rotation of lines in space, and are easy to implement. This section concludes with an example of a transformation from a solution in terms of TE_z and TM_z modes to one in terms of TE_x modes only, which also serves as additional validation of the MM formulation derived in section 3.2. The derivation of transformation matrices that can be applied to the GSMs to convert them freely from one modal representation to another is an original contribution to the field [Vale 2000b]. Although the transformation matrices apply to the GSM formulation directly, they can be modified to apply to other network parameters if so desired.

3.7.1 Relations between the modal sets

The various modal solution sets can be generated by solving the Helmholtz equations in terms of either the *x*, *y* or *z* oriented electric and magnetic field components [Harrington 1961]. This leads to the two solutions in eqn (3.38), with $\eta \in (x, y, z)$, $A_{\eta i}$ a normalisation constant, f a spatial distribution in *(x,y)*, k_x and k_y the spatial constants, γ the propagation constant which can be either real or imaginary for propagating and non-propagating modes respectively, i the mode number and z the direction of propagation.. All the other TM_η and TE_η field

components are derived from these two functions. Eqns. (3.2) can be seen to be a specific example of this general equation.

$$\begin{aligned} e_{\eta i}^{TM\eta} &= A_{\eta i}^{TM\eta} f_{e\eta}(k_{xi}x, k_{yi}y)\exp(-\gamma_i z) \\ h_{\eta i}^{TE\eta} &= A_{\eta i}^{TE\eta} f_{h\eta}(k_{xi}x, k_{yi}y)\exp(-\gamma_i z) \end{aligned} \tag{3.38}$$

The normalisation constants are chosen to ensure unity transported complex power, as defined in eqn (3.9), with $\bar{e}_{ti}$ and $\bar{h}_{ti}$ the normalised field components of the *i*-th mode, perpendicular to the direction of propagation.

By using orthogonality between modes of the same set and linear dependence of the respective solution sets, it can be shown that any TE^{η}_i or TM^{η}_i mode of a specific set, η, can be expressed as a weighted sum of only the TE^{ζ}_i and TM^{ζ}_i modes of any other set, ζ. This stands to reason, as the spatial field distributions of any given field component is dependant on the same $f(k_{xi}x, k_{yi}y)$ function irrespective of the mode set in which it is expressed. This leads to the general equality in eqn (3.40), where {} implies two equations.

$$\begin{bmatrix} \{\bar{e}_i^{TE\eta}, \bar{h}_i^{TE\eta}\} \\ \{\bar{e}_i^{TM\eta}, \bar{h}_i^{TM\eta}\} \end{bmatrix} = \left[M_i^{\eta\zeta}\right] \begin{bmatrix} \{\bar{e}_i^{TE\zeta}, \bar{h}_i^{TE\zeta}\} \\ \{\bar{e}_i^{TM\zeta}, \bar{h}_i^{TM\zeta}\} \end{bmatrix} \tag{3.39}$$

The transformation matrices $[M_i]$ are determined by equating the individual field magnitudes for each mode number, *i*, and solving the resulting two simultaneous equations, and are shown in eqn (3.40).

$$\begin{aligned} [M_i^{xy}] = [M_i^{yx}]^{-1} &= \frac{1}{\sqrt{k^2 - k_{xi}^2}\sqrt{k^2 - k_{yi}^2}} \begin{bmatrix} -k_{xi}k_{yi} & j\gamma_i k \\ -j\gamma_i k & -k_{xi}k_{yi} \end{bmatrix} \\ [M_i^{xz}] = [M_i^{zx}]^{-1} &= \frac{1}{k_{ci}\sqrt{k^2 - k_{xi}^2}} \begin{bmatrix} -\gamma_i k_{xi} & -jkk_{yi} \\ -jkk_{yi} & \gamma_i k_{xi} \end{bmatrix} \\ [M_i^{yz}] = [M_i^{zy}]^{-1} &= \frac{1}{k_{ci}\sqrt{k^2 - k_{yi}^2}} \begin{bmatrix} -\gamma_i k_{yi} & jkk_{xi} \\ jkk_{xi} & \gamma_i k_{yi} \end{bmatrix} \end{aligned} \tag{3.40}$$

The complete field vector of the i-th mode is obtained by multiplying the normalised field vectors with amplitude constants, a_i. If separate sets of constants are defined for the different mode sets, eqn (3.41) results for the E_t field, expressed in terms of two sets of modes. Using eqns (3.39) and (3.40) in eqn (3.41), it can easily be shown that eqn (3.42) holds for any two mode sets.

$$\begin{aligned}\bar{E}_i &= a_i^{TE\eta}\bar{e}_i^{TE\eta} + a_i^{TM\eta}\bar{e}_i^{TM\eta} \\ &= a_i^{TE\zeta}\bar{e}_i^{TE\zeta} + a_i^{TM\zeta}\bar{e}_i^{TM\zeta}\end{aligned} \tag{3.41}$$

$$\begin{bmatrix} a_i^{TE\eta} \\ a_i^{TM\eta} \end{bmatrix} = \left[M_i^{\eta\zeta}\right] \begin{bmatrix} a_i^{TE\zeta} \\ a_i^{TM\zeta} \end{bmatrix} \tag{3.42}$$

The transformation matrices above are two-by-two matrices designed to convert only one mode. Note that both the TE_ζ and TM_ζ modes are required for conversion. In cases where a TE or TM mode might not exist, the respective elements of the transform matrices will be zero in any case and the mode need not be specified.

3.7.2 Conversion between GSMs

The generalised scattering matrix for a physical two-port discontinuity can be written as in eqn (3.43), with the vector entries as defined in section 3.2.

$$\begin{bmatrix} \bar{b}_A^{TE} \\ \bar{b}_A^{TM} \\ \bar{b}_B^{TE} \\ \bar{b}_B^{TM} \end{bmatrix} = \begin{bmatrix} \left[S_{AA}^{TE-TE}\right] & \left[S_{AA}^{TM-TE}\right] & \left[S_{AB}^{TE-TE}\right] & \left[S_{AB}^{TM-TE}\right] \\ \left[S_{AA}^{TE-TM}\right] & \left[S_{AA}^{TM-TM}\right] & \left[S_{AB}^{TE-TM}\right] & \left[S_{AB}^{TM-TM}\right] \\ \left[S_{BA}^{TE-TE}\right] & \left[S_{BA}^{TM-TE}\right] & \left[S_{BB}^{TE-TE}\right] & \left[S_{BB}^{TM-TE}\right] \\ \left[S_{BA}^{TE-TM}\right] & \left[S_{BA}^{TM-TM}\right] & \left[S_{BB}^{TE-TM}\right] & \left[S_{BB}^{TM-TM}\right] \end{bmatrix} \begin{bmatrix} \bar{a}_A^{TE} \\ \bar{a}_A^{TM} \\ \bar{a}_B^{TE} \\ \bar{a}_B^{TM} \end{bmatrix} \tag{3.43}$$

The $\bar{a}$ and $\bar{b}$ vectors and the scattering matrix in eqn (3.43) can be regrouped as shown in eqn (3.44), where the subscript *reg* indicates the regrouped matrices.

$$\begin{bmatrix}\bar{b}_A\\ \bar{b}_B\end{bmatrix}_{reg}=\begin{bmatrix}\begin{bmatrix}b^{TE}_{A(1)}\\ b^{TM}_{A(1)}\end{bmatrix}\\ \begin{bmatrix}b^{TE}_{A(2)}\\ b^{TM}_{A(2)}\end{bmatrix}\\ \vdots\\ \begin{bmatrix}b^{TE}_{B(1)}\\ b^{TM}_{B(1)}\end{bmatrix}\\ \begin{bmatrix}b^{TE}_{B(2)}\\ b^{TM}_{B(2)}\end{bmatrix}\\ \vdots\end{bmatrix}\qquad\begin{bmatrix}\bar{a}_A\\ \bar{a}_B\end{bmatrix}_{reg}=\begin{bmatrix}\begin{bmatrix}a^{TE}_{A(1)}\\ a^{TM}_{A(1)}\end{bmatrix}\\ \begin{bmatrix}a^{TE}_{A(2)}\\ a^{TM}_{A(2)}\end{bmatrix}\\ \vdots\\ \begin{bmatrix}a^{TE}_{B(1)}\\ a^{TM}_{B(1)}\end{bmatrix}\\ \begin{bmatrix}a^{TE}_{B(2)}\\ a^{TM}_{B(2)}\end{bmatrix}\\ \vdots\end{bmatrix}\qquad\begin{bmatrix}\bar{b}_A\\ \bar{b}_B\end{bmatrix}_{reg}=[S_{reg}]\begin{bmatrix}\bar{a}_A\\ \bar{a}_B\end{bmatrix}_{reg}\tag{3.44}$$

Applying eqn (3.42) to the regrouped matrices, transformation of the scattering parameters from one mode set to another can easily be performed, as shown in eqn (3.45). Here, the subscript *{A,B}* indicates either port A or B, and *(i)* the mode number and the transformation matrices are as shown in eqn. (3.47). From eqn. (3.46) the GSM can be identified and is expressed in eqn. (3.48)

$$\begin{bmatrix}\bar{b}_A\\ \bar{b}_B\end{bmatrix}^{\eta}_{reg}=\begin{bmatrix}[\tilde{M}^{\eta\zeta}_A] & 0\\ 0 & [\tilde{M}^{\eta\zeta}_B]\end{bmatrix}\begin{bmatrix}\bar{b}_A\\ \bar{b}_B\end{bmatrix}^{\zeta}_{reg}=\begin{bmatrix}[\tilde{M}^{\eta\zeta}_A] & 0\\ 0 & [\tilde{M}^{\eta\zeta}_B]\end{bmatrix}[S^{\zeta}_{reg}]\begin{bmatrix}[\tilde{M}^{\zeta\eta}_A] & 0\\ 0 & [\tilde{M}^{\zeta\eta}_B]\end{bmatrix}\begin{bmatrix}\bar{a}_A\\ \bar{a}_B\end{bmatrix}^{\eta}_{reg}\tag{3.46}$$

$$[\tilde{M}^{\eta\zeta}_{\{A,B\}}]=\begin{bmatrix}[M^{\eta\zeta}_{\{A,B\}(1)}] & 0 & 0 & 0\\ 0 & [M^{\eta\zeta}_{\{A,B\}(2)}] & 0 & \vdots\\ \vdots & 0 & \ddots & 0\\ 0 & 0 & 0 & [M^{\eta\zeta}_{\{A,B\}(N)}]\end{bmatrix}\tag{3.47}$$

$$[S^{\eta}_{reg}]=\begin{bmatrix}[\tilde{M}^{\eta\zeta}_A] & 0\\ 0 & [\tilde{M}^{\eta\zeta}_B]\end{bmatrix}[S^{\zeta}_{reg}]\begin{bmatrix}[\tilde{M}^{\zeta\eta}_A] & 0\\ 0 & [\tilde{M}^{\zeta\eta}_B]\end{bmatrix}\tag{3.48}$$

By performing an inverse regrouping of the scattering matrix, the final scattering matrix is obtained.

3.7.3. Example of conversion in a mode matching application

As an example of the use of these transformations, an E-plane step in rectangular waveguide is solved by the mode-matching technique in terms of TE_z/TM_z modes. The resulting GSM is then converted to one in terms of TE_x/TM_x modes. The validity of the transformation is illustrated by solving the same waveguide problem in terms of TE_x modes with an independent mode matching code, and comparing the two results. Figure 3.25 shows various GSM entries for the two solutions as a function of frequency. In the limit of an infinite number of modes in both mode-matching solutions, the results could be expected to agree exactly. Due to limiting the number of modes, however, a small discrepancy can be observed. As is expected, the conversion in the example above correctly predicted zero cross-coupling between TE_x and TM_x modes incident on this type of discontinuity. Tests on a single-plane step in the horizontal direction correctly showed zero cross-coupling between TE_y and TM_y modes, once again verifying the conversion process.

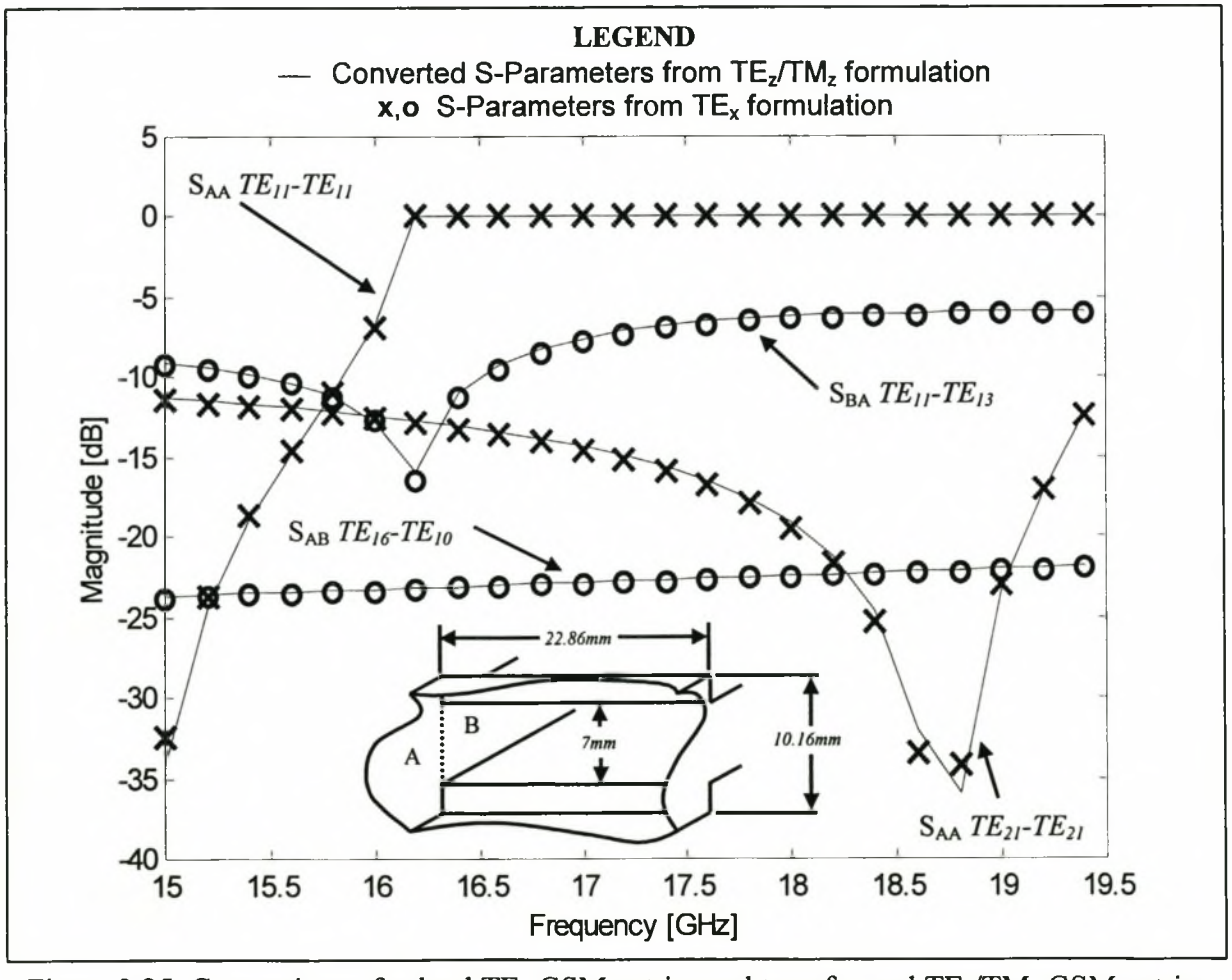


Figure 3.25. Comparison of solved TE_x GSM entries and transformed TE_z/TM_z GSM entries

3.8 Conclusion

This chapter has detailed the development of analysis techniques for the type of design problems to be addressed by the CAD techniques presented in later chapters.

First, the mode matching formulation was derived for steps in rectangular waveguide. The generalized scattering matrix formulation was then adopted for use in cascading these steps, allowing the analysis of general waveguide structures. For the purposes of CAD, however, such analysis must be fast.

The following sections detailed the enhancement of the basic mode matching technique by the introduction of a new equivalent circuit formulation for mode matching. This equivalent circuit was exploited to aid in the intelligent selection of modes for the mode matching operation. A novel technique was described which allows the mode matching problem to be broken into a series of independent sub-problems which can be separately solved and cascaded using the GSM formulation. This new analysis technique yielded a massive increase in analysis speed over the conventional approach with no loss in flexibility.

In order to push even higher analysis speeds, the concept of surrogate models was briefly investigated. A robust, multivariate, multiple-output interpolation scheme was proposed as a possible surrogate for EM analysis in the context of the CAD procedure to be pursued by this investigation. The model was successfully tested on the analysis of a band pass filter, showing, as expected, a large increase in analysis speed, coupled to a drop in accuracy.

Finally, an original method of converting scattering parameters expressed in terms of different modal sets was presented. This useful development will prove important in the next chapter, where novel measurement techniques are developed for overmoded structures in waveguide.

Chapter 4

Measurement Technique Development

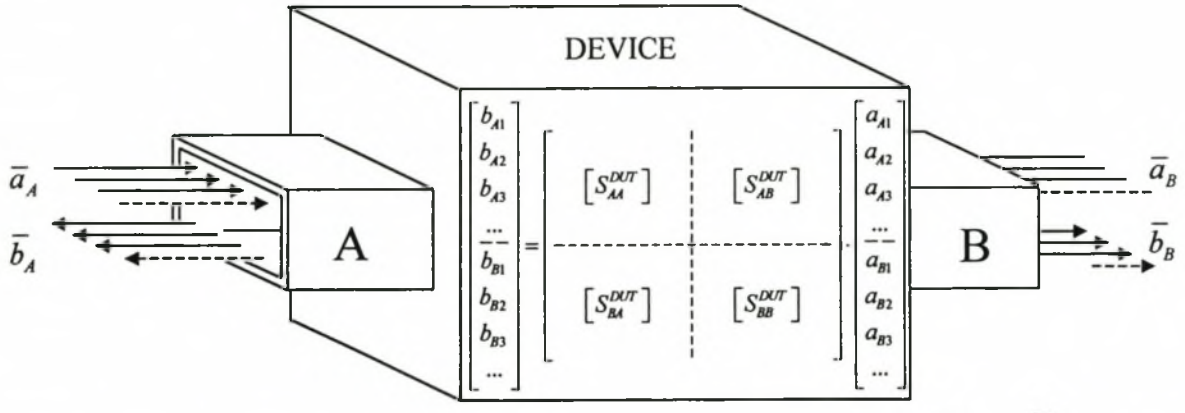


Figure 4.1. A multimode device and its GSM representation. The vectors $\bar{a}$ and $\bar{b}$ contain the amplitudes of the incident and reflected modes at each port respectively.

4.1 Introduction

A critical component of design is the ability to verify the prototype structure by measurement. Since this dissertation concerns itself with the development of complex multimode structures in waveguide, it is essential to establish a measurement scheme that can provide meaningful feedback on the performance of the devices. Measurement of devices supporting multiple propagating modes at the ports, such as the one depicted in figure 4.1, is usually a difficult and time-consuming operation - difficult because the contributions of the different modes have to be separated by using measurements at one physical port, and time consuming because each mode constitutes a different electrical port, creating a multi-port device which can only be characterized by a generalized scattering matrix. Typical examples of such devices include chokes for microwave heating applications [Vale 2000d] and mode converters [Buckley 1990, Lunéville 1998].

Various approaches have been proposed in the literature to deal with this type of problem. The earliest techniques, described by Beck, relied on the use of mode couplers or, alternatively, very

long lengths of line to exploit the difference in propagation speed of the modes [Beck 1955]. Felsen et al. describe a technique using a plunger to vary the resonance condition of a cylindrical guide cavity for various modes [Felsen 1959]. Price suggests the use of probes to sample field strengths along the guide walls and the use of Fourier techniques to extract the modal content [Price 1959]. This approach is expanded upon by others to make it more accurate and easier to use [Sharp 1961, Levinson 1966, Taub 1962]. In particular, Levinson et al. detail efforts made to determine the relative power distribution in the propagating modes in waveguide, and propose the use of an enlarged waveguide incorporating a probe section to sample the fields along the waveguide walls, using a phase-sensitive detector is to discriminate between the various modes. Later, Carmel et al. propose the use of a liquid crystal sheet to identify modal patterns in overmoded waveguide [Carmel 1984]. Most recently, Sequinot et al., present a method for multiport device measurement that makes use of multimode-to-*N*-line transitions in microstrip [Sequinot 1998]. With the exception of the last method, none of the methods discussed above offer accurate broadband measurement. Sequinot's method, while by far the most accurate, requires the characterisation of a *2N* port with a network analyzer. For five propagating modes at each of the two ports, this requires more than 200 measurements, which limits the usability significantly.

This chapter introduces two novel techniques, both using a significantly reduced number of measurements, for the accurate determination of the generalized S-matrix of two-port networks which exhibit multiple propagating modes at each port. In each case the reduction in complexity is achieved by the use of lightly coupled co-axial probe feeds as transitions from the overmoded guide to a single-mode co-axial line, in conjunction with alternative waveguide mode sets. The use of light coupling in the transitions is markedly different from the normal case, and when described in terms of alternative mode sets, produces a single-to-multimode transition with a very simple scattering matrix. The first technique uses such transitions, both terminated with a matched load on one side, and inserts different lengths of uniform waveguide between the transitions and the device. This allows the determination of the S-Matrix by solving sets of simultaneous linear equations. The second technique terminates one or both of the transitions with a sliding short, and requires the solution of a complex set of non-linear equations. The advantage of the latter method is that it eliminates the effort of changing waveguide sections, making it easier to obtain an over-determined set of equations, which can be used to reduce the effects of measurement noise and

error. Each technique is presented in general form, which accounts for multimode devices incorporating cross-coupling. Where there is limited cross-coupling between modes, considerable simplifications can be made and these are also presented where applicable.

The chapter starts by addressing the aspects of mode excitation with lightly coupled E- or H-field probes and the single-to-multimode transitions constructed from them. This is followed by calibration procedures for these transitions, and the two measurement procedures. The chapter concludes by comparing results of the two measurement procedures with the predicted response of an overmoded bandstop filter, whose design will be detailed in the next chapter. Accuracies to within a few dB from predicted results are obtained for scattering parameters as low as -50dB.

4.2 Co-axial to Multimode Waveguide Transitions

To measure a multimode device, some sort of controllable mode excitation is required at the input and output ports. This is achieved here with a lightly coupled single-mode co-axial probe, feeding into an overmoded rectangular waveguide. The requirement of light coupling is of key importance, as this makes the proposed reduction in the number of measurements possible. In modelling the transition, we will consider a probe section intruding into open waveguide to create a 3-(physical) port device, as shown in figure 4.2.

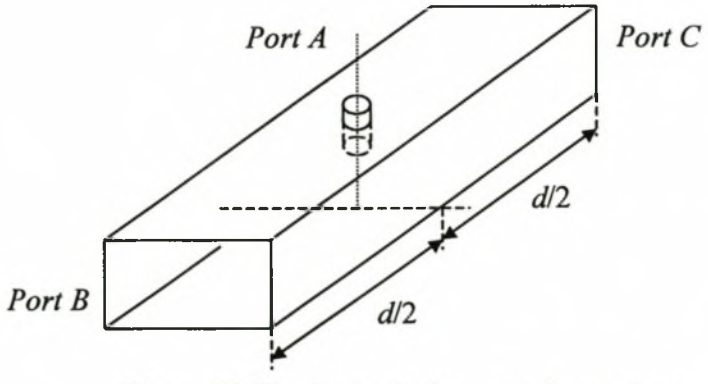


Figure 4.2 The 3-physical-port probe section

This transition is described by the generalized scattering matrix in eqn. (4.1). This is a subset of the generalised scattering matrix of the transition, consisting of only the propagating modes which couple to the probe, and implies that the reference planes at the ports are positioned in such a way that all non-propagating modes have died out. Note that all waveguide modes not excited by the probe, are excluded. In order to obtain a description containing all the propagating modes, it may, therefore, be necessary to use a few different probe placements.

$$\begin{bmatrix} b_A \\ \overline{b}_B \\ \overline{b}_C \end{bmatrix} = \begin{bmatrix} S_{AA} & \overline{S}_{AB} & \overline{S}_{AC} \\ \overline{S}_{BA} & [S_{BB}] & [S_{BC}] \\ \overline{S}_{CA} & [S_{CB}] & [S_{CC}] \end{bmatrix} \begin{bmatrix} a_A \\ \overline{a}_B \\ \overline{a}_C \end{bmatrix} \tag{4.1}$$

Although only a subset of the full matrix, the characterisation of all the unknowns in equation (4.1) is still a formidable task, requiring a large number of calibration measurements. To alleviate this problem, two mechanisms are proposed: (i) a probe extending only a very short distance into the waveguide, i.e. a lightly coupled probe, and (ii) to model the waveguide in terms of TE_y/TM_y and/or TE_x/TM_x mode sets, depending on the probe orientation.

By implementing (i), a transition is formed with the following characteristics:

- negligible probe-induced cross-coupling between waveguide modes
- negligible probe-induced reflection of waveguide modes incident from ports B or C.
- vertically (horizontally) orientated probes excite electric fields with negligible horizontal (vertical) components.

This has a huge effect on the model, as replacement of all the negligible quantities by zero causes $[S_{BB}]$ and $[S_{CC}]$ in equation (4.1) to become zero, while $[S_{CB}]$ and $[S_{CC}]$ both reduce to the diagonal matrix $\left[\widetilde{D}\right]$ in equation (4.2), with β_n representing the propagation constants of the propagating modes, $n = 1..N$, in the waveguide.

$$[S_{BC}] = [S_{CB}] = \left[\widetilde{D}\right] = diag(e^{-j\beta_n d}) \tag{4.2}$$

Regarding (ii), it will be shown that it is undesirable to work with degenerate modes (modes with different field distributions but equal propagation constants, e.g. TE_{z11} and TM_{z11}) as they can introduce singular behaviour in both the calibration and device measurement procedures. This is very difficult to avoid when working with TE_z/TM_z mode sets. However, working instead with the alternative TE_y/TM_y and/or TE_x/TM_x mode sets can readily solve this problem, as a *y*-directed probe which extends only a very short distance into the waveguide, will excite a negligible E_x-field, and therefore only TE_x and no TM_x modes. Conversely, a probe exciting a negligible E_y-field, will only excite TE_y and no TM_y fields. It should be evident that this approach has the added benefit that, if only TE modes can be used, the number of modes needed to describe the problem is halved, reducing the complexity of the matrix significantly. Using these alternative mode sets places no limitation on the technique, as S-parameter results expressed in terms of one mode set can be easily transformed to other mode sets by using the transformation matrices derived in chapter 3.

Finally, due to reciprocity and symmetry in the *z*-direction, we have $\bar{S}_{CA} = \bar{S}_{BA} = \bar{S}'_{AC} = \bar{S}'_{AB}$, resulting in the final description of the transition in terms of only the coupled, non-degenerate modes, shown in equation (4.3). This matrix contains only $N+1$ complex unknowns, i.e. S_{AA} and $\bar{S}_{BA}$, with N the number of coupled modes in the waveguide. It is clear that this characterisation of the transition is an approximate one, which will limit the accuracy and dynamic range of the technique. Nevertheless, for probe couplings of the order of about -10dB to desired modes, very accurate results have been obtained. Coupling through narrow slots can also be used to good effect instead of (or in addition to) probes, as the amount of unwanted field coupling can be better controlled. This approach has the drawback of slightly more complicated manufacture.

$$\begin{bmatrix} b_A \\ \bar{b}_B \\ \bar{b}_C \end{bmatrix} = \begin{bmatrix} S_{AA} & \bar{S}'_{BA} & \bar{S}'_{BA} \\ \bar{S}_{BA} & [0] & [\tilde{D}] \\ \bar{S}_{BA} & [\tilde{D}] & [0] \end{bmatrix} \begin{bmatrix} a_A \\ \bar{a}_B \\ \bar{a}_C \end{bmatrix} \tag{4.3}$$

4.3 Calibration

Before the device can be measured, the transition(s) needs to be characterised. Three calibration techniques are presented to achieve this, discussed in order of decreasing measurement effort and increasing numerical effort. They are a standard load/offset-short calibration, a back-to-back calibration of two identical transitions with the free ports terminated by ideal loads, and a back-to-back calibration of two identical transitions with one free port terminated by an ideal load and the other by a sliding short. Note that in all cases measured parameters will be denoted as T for transmission measurements and Γ for reflection measurements. The three cases are summarized in figure 4.3.

4.3.1 Standard Calibration

In this case, shown in figure 4.3(a), one of the ports of the transition (in this case port C) is terminated in an ideal multimode load. The resulting transition is described by eqn (4.4).

$$\begin{bmatrix} b_A \\ \bar{b}_B \end{bmatrix} = \begin{bmatrix} S_{AA} & \bar{S}'_{BA} \\ \bar{S}_{BA} & [0] \end{bmatrix} \begin{bmatrix} a_A \\ \bar{a}_B \end{bmatrix} \tag{4.4}$$

To obtain S_{AA}, port B is simply terminated with an ideal multimode load, making the measured reflection from port A, $\Gamma_{11}(load)$, equal to S_{AA}. If next port B is terminated in an ideal sliding offset short of length $l_m/2$, the measured reflection at port A will be as in eqns. (4.5).

$$\begin{gathered} \Gamma_{11}(l_m) = S_{AA} - \bar{S}_{BA}{}^{t}\left[\tilde{L}_m\right]\bar{S}_{BA} = S_{AA} - \bar{L}_m \cdot \bar{S}_{BA}^2 \\ with \\ \bar{L}_m = \exp\left(-j\left[\beta_1 l_m \quad \ldots \quad \beta_N l_m\right]\right) \quad and \quad \left[\tilde{L}_m\right] = diag(\bar{L}_m) \end{gathered} \tag{4.5}$$

If this is repeated K times, for K different lengths of line, a full set of linear equations are obtained for the determination of S_{BA} as shown in eqns. (4.6)

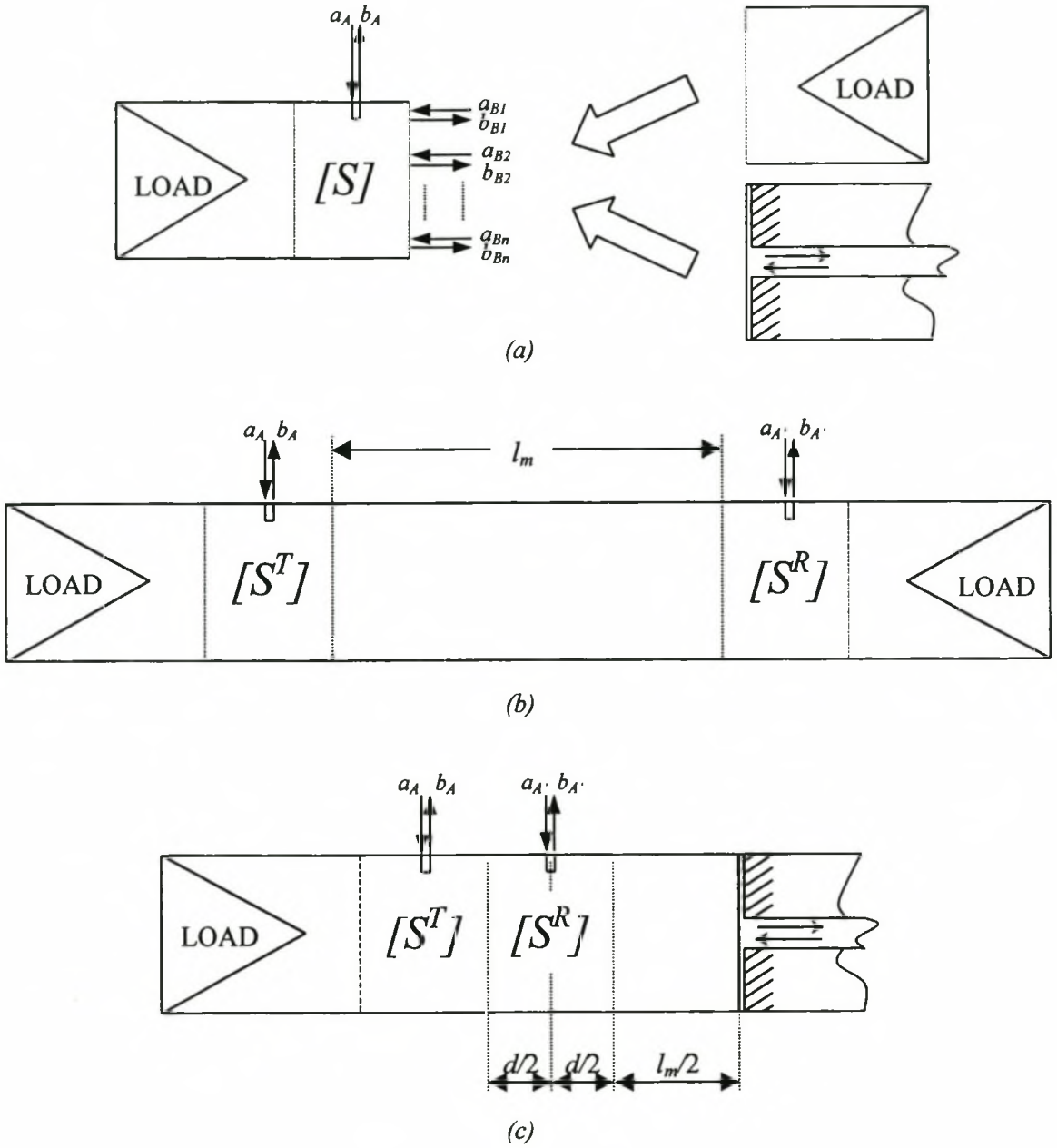


Figure 4.3. *(a)* Standard load/offset short calibration. *(b)* Back to back calibration of two identical transitions separated by line sections with the free ports terminated by ideal loads. *(c)* Back to back calibration of two identical transitions with one free port terminated by an ideal load and the other by a sliding short.

$$\bar{\Gamma}_{11} - S_{AA} = [E] \cdot \bar{S}_{BA}^2$$
$$\textit{with} \tag{4.6}$$
$$\bar{\Gamma}_{11} = \left[\Gamma_{11}(l_1) \quad \Gamma_{11}(l_2) \quad \ldots \quad \Gamma_{11}(l_K)\right]^t \quad \textit{and} \quad [E] = \left[\bar{L}_1^t \quad \bar{L}_2^t \quad \ldots \quad \bar{L}_K^t\right]^T$$

If K is greater than or equal to the number of propagating modes, N, and $[E]$ is not singular, this system can be solved for the unknown parameters, $\bar{S}_{BA}^2$. Note that for $[E]$ to be non-singular, it is necessary that no two or more elements of the $\bar{L}_{1..K}$ vectors be the same, prohibiting the use of lines which differ by multiples of half wavelengths, and the occurrence of degenerate modes. The first of these conditions is easy to satisfy, and the second is inherently satisfied by virtue of the choice of mode set. The use of $K > N$ measurements gives an over-determined system, which can be exploited to reduce the effects of measurement noise and error by using an iterative solver (such as least squares). This process must be repeated for both transitions.

4.3.2 Reduced Calibration 1

An alternative TRL-type calibration method, shown in figure 4.3(b), also terminates one port of each transition in a perfect load, and uses a number of back to back measurements of the transitions, separated by varying lengths of lines. A section of overmoded uniform guide of length l_m , inserted between the transitions, results in a measured transmission, $T_{21}(l_m)$, as in eqn. (4.7), where $\left(\overline{S_{BA}^R S_{BA}^T}\right)$ implies a vector whose elements are products of piecewise multiplied elements of the $\bar{S}_{BA}^R$ and $\bar{S}_{BA}^T$ vectors.

$$T_{21}(l_m) = \left(\bar{S}_{BA}^R\right)^t \left[\tilde{L}_m\right] \bar{S}_{BA}^T = \bar{L}_m \cdot \left(\overline{S_{BA}^R S_{BA}^T}\right) \tag{4.7}$$

If this is repeated K times for different lengths, l_m, as before, a set of linear equations can be written as in eqn. (4.8),

$$\bar{T}_{21} = [E]\left(\overline{S_{BA}^R S_{BA}^T}\right)$$
$$\textit{with} \quad \bar{T}_{21} = \left[T_{21}(l_1) \quad T_{21}(l_2) \quad \ldots \quad T_{21}(l_K)\right]^T \tag{4.8}$$
$$\textit{and} \quad [E] \quad \textit{defined as before}$$

As before, the system is solvable for $K \geq N$, provided $[E]$ is not singular. The S_{AA}^{T} and S_{AA}^{R} parameters can be found from the mean measured values $\Gamma_{11}(l_m)$ and $\Gamma_{22}(l_m)$ over the various line lengths, or by termination in perfect loads, as in the standard calibration (section 4.3.1)

This approach provides a reduction in the number of measurements needed as compared to the standard calibration, but is not able to uniquely identify the parameters of each transition separately, unless the transitions R and T are assumed to be identical. It will, however be shown that device de-embedding can in some cases be successfully accomplished with the transition parameters expressed in this product form. Note that this approach requires that the waveguide dimensions on sides T and R are the same.

4.3.3 Reduced Calibration 2

The introduction of a number of line lengths between two transitions can be very tedious, especially if a large number of modes must be measured. This third technique achieves the same number of independent measurements as the previous technique, but alleviates the need for constant replacement of lines by making use of a sliding short to terminate one of the transitions.

Using the setup in figure 4.3(c), the measured transmission for a given offset length is given in eqn (4.9) which, with an extended L vector defined as in eqn. (4.10), can be rewritten as eqn. (4.11).

$$\mathrm{T}_{21}(l_m) = \left(\bar{S}_{BA}^{R}\right)^{t} \cdot \left([I] - [\tilde{L}_m][\tilde{D}]\right) \cdot \bar{S}_{BA}^{T} \tag{4.9}$$

$$\bar{L}_m^e = \left[1 - \exp\{-j\beta_1(l_m + d)\} \quad 1 - \exp\{-j\beta_2(l_m + d)\} \quad \ldots \quad 1 - \exp\{-j\beta_N(l_m + d)\}\right] \tag{4.10}$$

$$\mathrm{T}_{21}(l_m) = \bar{L}_m^e \cdot \left(\overline{S_{BA}^{R} S_{BA}^{T}}\right) \tag{4.11}$$

If K measurements are performed with K offset lengths, the system of linear equations can be expressed as in eqn. (4.12).

$$\bar{\mathrm{T}}_{21} = \left[E^e\right] \cdot \left(\overline{S_{BA}^{R} S_{BA}^{T}}\right) \tag{4.12}$$

The conditions for $[E^e]$ being non-singular are as before, with, once again, the option of exploiting an over-determined system for better accuracy. Note that, like the previous calibration technique, this approach is only suitable when measuring with the same size waveguide on sides T and R, and only yields the S_{BA} terms in product form. S_{AA} can again be calculated by using the standard calibration method.

4.4 Device Measurement

Once the transitions are characterised accurately, the device itself can be measured. Two techniques will be described, again in order of decreasing measurement effort and increasing numerical effort. As with the calibration, the first technique uses transitions terminated at the ends by ideal loads, while the second terminates the transitions with sliding shorts.

4.4.1 Transitions terminated with matched loads

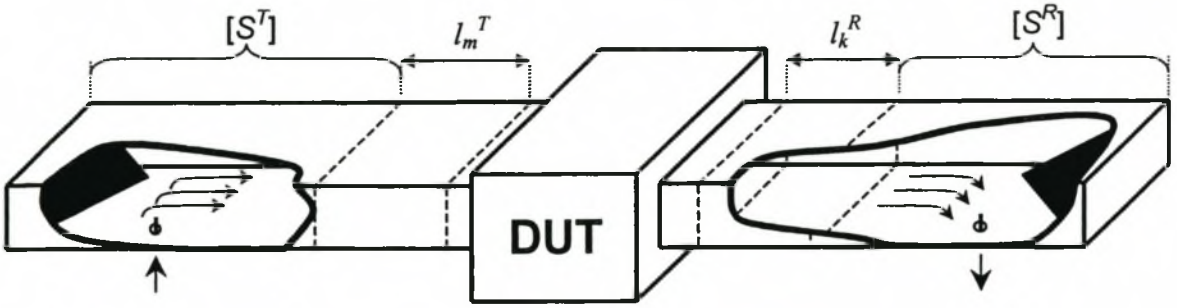


Figure 4.4. Measurement setup with load-terminated transitions

This technique uses variable line sections, l_m^T and l_k^R, inserted between the device ports and the transitions, both of which are terminated at the outside ports with ideal loads. With this configuration, energy reflected back from the DUT is predominantly absorbed, while energy which propagates through the DUT is sampled at the receiving transition's probe, and the predominant remaining energy is absorbed by the load terminating the receive transition, causing negligible energy to be reflected back to the DUT. For various choices of line lengths, the different

propagation constants of the propagating modes will cause the modes to recombine differently at the receiving transition's probe. If enough length combinations are used, the characteristics of the device can be extracted from the measured transmission coefficients of the system.

In the same way as before, we define,

$$\begin{aligned} \bar{L}_m^{\,T} &= \exp\left(-j\left[\beta_1^{\,T} l_m^{\,T} \quad \beta_2^{\,T} l_m^{\,T} \quad \ldots \quad \beta_N^{\,T} l_m^{\,T}\right]\right) \\ \bar{L}_k^{\,R} &= \exp\left(-j\left[\beta_1^{\,R} l_k^{\,R} \quad \beta_2^{\,R} l_k^{\,R} \quad \ldots \quad \beta_Q^{\,R} l_k^{\,R}\right]\right) \\ \left[\tilde{L}\right] &= diag\left(\bar{L}\right) \end{aligned} \tag{4.13}$$

with N propagating modes in side T, and Q propagating modes in side R. Then, from cascading the transition and device parameters, the resultant measured parameters will be

$$\begin{aligned} \Gamma_{11}\left(l_m^{\,T}\right) &= S_{AA}^{T} + \bar{S}_{AB}^{T} \cdot \left[\tilde{L}_m^{T}\right] \cdot \left[S_{AA}^{DUT}\right] \cdot \left[\tilde{L}_m^{T}\right] \cdot \bar{S}_{BA}^{T} \\ \mathrm{T}_{12}\left(l_m^{\,T}, l_k^{\,R}\right) &= \bar{S}_{AB}^{T} \cdot \left[\tilde{L}_m^{T}\right] \cdot \left[S_{AB}^{DUT}\right] \cdot \left[\tilde{L}_k^{R}\right] \cdot \bar{S}_{BA}^{R} \\ \mathrm{T}_{21}\left(l_k^{\,R}, l_m^{\,T}\right) &= \bar{S}_{AB}^{R} \cdot \left[\tilde{L}_k^{R}\right] \cdot \left[S_{BA}^{DUT}\right] \cdot \left[\tilde{L}_m^{T}\right] \cdot \bar{S}_{BA}^{T} \\ \Gamma_{22}\left(l_k^{\,R}\right) &= S_{AA}^{R} + \bar{S}_{AB}^{R} \cdot \left[\tilde{L}_k^{R}\right] \cdot \left[S_{BB}^{DUT}\right] \cdot \left[\tilde{L}_k^{R}\right] \cdot \bar{S}_{BA}^{R} \end{aligned} \tag{4.14}$$

Now, since for two column vectors, $\bar{A}$ and $\bar{B}$, eqn. (4.15) holds, the measured transmission can be written as in eqn. (4.16)

$$\bar{A}^{t} \cdot diag\left(\bar{B}\right) = \bar{B}^{t} \cdot diag\left(\bar{A}\right) \tag{4.15}$$

$$\mathrm{T}_{21}\left(l_k^{\,R}, l_m^{\,T}\right) = \bar{L}_k^{\,R} \cdot diag\left(\bar{S}_{AB}^{R}\right) \cdot \left[S_{BA}^{DUT}\right] \cdot diag\left(\bar{S}_{BA}^{T}\right) \cdot \left(\bar{L}_m^{\,T}\right)^{t} \tag{4.16}$$

When using line lengths, $l_k^{\,R}$ and $l_m^{\,T}$, with $k = 1..K$ and $m = 1..M$, the resulting system of linear equations can be written as eqn. (4.17).

$$\begin{aligned} &\left[\mathrm{T}_{21}\right] = \left[E^{R}\right] \cdot \left[\tilde{S}_{AB}^{R}\right] \cdot \left[S_{BA}^{DUT}\right] \cdot \left[\tilde{S}_{BA}^{T}\right] \cdot \left[E^{T}\right]^{t} \\ &with \quad \mathrm{T}_{21}(k,m) = \mathrm{T}_{21}\left(l_k^{\,R}, l_m^{\,T}\right) \end{aligned} \tag{4.17}$$

From this, $\left[S_{BA}^{DUT}\right]$ can be calculated by either having K=Q and M=N and solving the set of equations, or by solving the over-determined system (K>Q and/or M>N) to enhance resistance to noise and measurement error. This, however, involves more measurements.

The determination of $S_{AA}{}^{DUT}$ and $S_{BB}{}^{DUT}$ using this procedure is somewhat more complicated than with the transmission parameters, as a full set of non-linear equations, each one containing all the elements of $S_{AA}{}^{DUT}$ or $S_{BB}{}^{DUT}$ is required. If these parameters are required, the method in the following section should be preferred. It should be clear that this procedure may not yield a full description of the device, as modes which are not excited by the probes will not be included. This is often the case, as degenerate modes are excluded by design, owing to their undesirable effects on the [E] matrices. To get a full description of the device, including all propagating modes, different probe positions and orientations should be used.

An interesting special case often occurring in waveguide, is that of a device exhibiting cross-coupling between degenerate modes only. As discussed before, any measurement set of such a device will yield a subset of the full scattering parameters, containing no degenerate modes, and as no other cross-coupling exists, these will consist of purely diagonal matrices. In this case, one of the variable lengths, $l_m{}^T$ or $l_k{}^R$, may be set permanently to zero, reducing eqn. (4.17) to the following linear system,

$$\overline{\mathrm{T}}_{21} = \left[E^{(R\ or\ T)}\right] \cdot \left[\widetilde{S}_{BA}^{RT}\right] \cdot \overline{S}_{BA}^{DUT} \quad with \quad \widetilde{S}_{BA}^{RT} = diag\left(\overline{S_{BA}^{R} S_{BA}^{T}}\right) \tag{4.18}$$

where $\overline{S}_{BA}^{DUT}$ is now a vector of the diagonal elements of the reduced $\left[S_{BA}^{DUT}\right]$matrix and $\overline{\mathrm{T}}_{21}$ is a vector of measured T_{21} for different lengths corresponding to the rows of [E]. It should be evident that this special case requires significantly fewer measurements in both the characterisation of the device as well as the transitions, as the $\left(\overline{S_{BA}^{R} S_{BA}^{T}}\right)$ term is only required in product form. In this case, the determination of $S_{AA}{}^{DUT}$ and $S_{BB}{}^{DUT}$ also reduces to a simple solution.

4.4.2 Transitions terminated with offset shorts

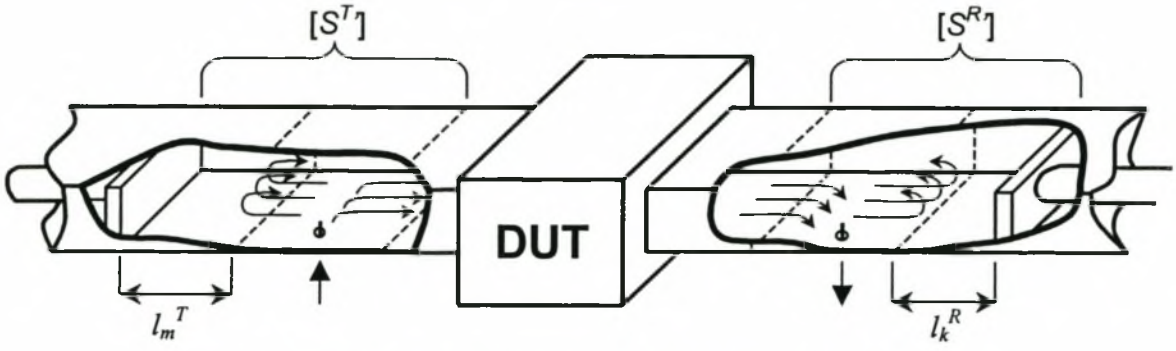


Figure 4.5. Measurement setup with the short-terminated transitions.

This method uses transitions both terminated at the open port with sliding offset shorts. A great advantage of this method is the ease with which measurements can be made for the various choices of lengths, l_k^R and l_m^T, requiring only the manipulation of the sliding shorts. A considerable improvement in measurement accuracy also results due to reduced mechanical disturbance of the measurement setup. The use of short circuit terminations provides an additional advantage in the form of a higher signal strength throughout the system, as energy is no longer absorbed by terminating loads (which might not be perfectly ideal). Finally, the full S^{DUT} matrix is obtained from one solution, as opposed to the previous method.

A fundamental disadvantage of this method is the substantial gain in complexity. In contrast to the previous method, where a subset of the device scattering matrix is determined directly, the use of offset shorts as terminations creates a new scattering matrix, that of the device between short circuits. Such a matrix normally contains non-linear combinations of the device scattering matrix, including those not excited by the probes. It follows that a sufficiently large number of measurements must be made to determine all the elements of the scattering matrix relating to propagating modes, and that a complex non-linear solver must be used to find the solution.

From the three-port description of the transition, the scattering parameters of the shorted transitions can be found as in eqns. (4.19) and (4.20), where port 'A' refers to the coaxial side and

port 'B' to the overmoded waveguide side of the transition in each case. The system parameters then follow in eqns. (4.21).

$$\begin{bmatrix} S_{AA}^{T}{}' & \bar{S}_{AB}^{T}{}' \\ \bar{S}_{BA}^{T}{}' & [S_{BB}^{T}{}'] \end{bmatrix} = \begin{bmatrix} S_{AA}^{T} - \bar{S}_{AC}^{T} \cdot [\tilde{L}_m{}^{T}] \cdot \bar{S}_{CA}^{T} & \bar{S}_{AB}^{T} - \bar{S}_{AC}^{T} \cdot [\tilde{L}_m{}^{T}] \cdot [S_{CB}^{T}] \\ \bar{S}_{BA}^{T} - [S_{BC}^{T}] \cdot [\tilde{L}_m{}^{T}] \cdot \bar{S}_{CA}^{T} & -[S_{BC}^{T}] \cdot [\tilde{L}_m{}^{T}] \cdot [S_{CB}^{T}] \end{bmatrix} \tag{4.19}$$

$$\begin{bmatrix} S_{AA}^{R}{}' & \bar{S}_{AB}^{R}{}' \\ \bar{S}_{BA}^{R}{}' & [S_{BB}^{R}{}'] \end{bmatrix} = \begin{bmatrix} S_{AA}^{R} - \bar{S}_{AC}^{R} \cdot [\tilde{L}_k{}^{R}] \cdot \bar{S}_{CA}^{R} & \bar{S}_{AB}^{R} - \bar{S}_{AC}^{R} \cdot [\tilde{L}_k{}^{R}] \cdot [S_{CB}^{R}] \\ \bar{S}_{BA}^{R} - [S_{BC}^{R}] \cdot [\tilde{L}_k{}^{R}] \cdot \bar{S}_{CA}^{R} & -[S_{BC}^{R}] \cdot [\tilde{L}_k{}^{R}] \cdot [S_{CB}^{R}] \end{bmatrix} \tag{4.20}$$

$$\begin{aligned} \Gamma_{11}(l_k{}^{R}, l_m{}^{T}) &= S_{AA}^{\Delta} + \bar{S}_{AB}^{\Delta}([I] - [S_{BB}^{R}{}'][S_{BB}^{\Delta}])^{-1}[S_{BB}^{R}{}'] \cdot \bar{S}_{BA}^{\Delta} \\ T_{12}(l_k{}^{R}, l_m{}^{T}) &= \bar{S}_{AB}^{\Delta}([I] - [S_{BB}^{R}{}'][S_{BB}^{\Delta}])^{-1} \bar{S}_{BA}^{R}{}' \\ T_{21}(l_k{}^{R}, l_m{}^{T}) &= \bar{S}_{AB}^{R}{}'([I] - [S_{BB}^{\Delta}][S_{BB}^{R}{}'])^{-1} \bar{S}_{BA}^{\Delta} \\ \Gamma_{22}(l_k{}^{R}, l_m{}^{T}) &= S_{AA}^{R}{}' + \bar{S}_{AB}^{R}{}'([I] - [S_{BB}^{\Delta}][S_{BB}^{R}{}'])^{-1}[S_{BB}^{\Delta}] \cdot \bar{S}_{BA}^{R}{}' \end{aligned} \tag{4.21}$$

$$\begin{aligned} \textit{with} \quad S_{AA}^{\Delta} &= S_{AA}^{T}{}' + \bar{S}_{AB}^{T}{}'([I] - [S_{AA}^{DUT}][S_{BB}^{T}{}'])^{-1}[S_{AA}^{DUT}] \cdot \bar{S}_{BA}^{T}{}' \\ \bar{S}_{AB}^{\Delta} &= \bar{S}_{AB}^{T}{}'([I] - [S_{AA}^{DUT}][S_{BB}^{T}{}'])^{-1}[S_{AB}^{DUT}] \\ \bar{S}_{BA}^{\Delta} &= [S_{BA}^{DUT}]([I] - [S_{BB}^{T}{}'][S_{AA}^{DUT}])^{-1} \bar{S}_{BA}^{T}{}' \\ [S_{BB}^{\Delta}] &= [S_{BB}^{DUT}] + [S_{BA}^{DUT}]([I] - [S_{BB}^{T}{}'][S_{AA}^{DUT}])^{-1}[S_{BB}^{T}{}'][S_{AB}^{DUT}] \end{aligned}$$

Using a minimum of $N \times Q$ measurements with different combinations of lengths, $l_k{}^{R}$ and $l_m{}^{T}$, each entry of the device scattering matrix can be identified uniquely. More length combinations and/or assuming symmetry of the $[S^{DUT}]$ matrices results in an over-determined system, typically speeding up the convergence of the iterative solver and improving the measurement accuracy. Note that the device parameters should be optimised to fit all the measured system parameters, Γ_{11}, Γ_{22}, T_{12} and T_{21} as closely as possible. There is, fortunately, no real penalty for using extra line sections, as this only requires manipulating the sliding short terminations.

As before, assumptions regarding the DUT can be used to make simplifications to the equations for the system parameters. If the same special case as before is considered, where the device

allows cross-coupling only to degenerate modes, then one of the sliding short terminations can be replaced with an ideal matched load. If this is done on the 'T' side, then the device $[S_{BA}^{DUT}]$and $[S_{BB}^{DUT}]$ matrices can be determined with a reduced set of non-linear equations. These equations can be easily found by substituting the parameters for a matched load, given in equation (4.4) for the [S^{T}'] parameters in equation (4.21) and using semi-diagonal matrices for the device parameters.

4.5 General Issues

The accuracy of the techniques described above depend largely on three considerations that must be taken into account during the measurement process. In addition to these considerations, it should be noted that the reliance on phase information inherent to the technique makes it critical that the physical measurement setup be disturbed as little as possible during measurements.

4.5.1 Choices of line lengths

The spread of values of lengths, l_m^T and l_k^R, used for calibration and device measurement can in many cases be chosen to improve tolerance to both noise and errors introduced by transition and device idealization. Monte Carlo analyses have shown that maximizing the determinant of the $[E]$ and $[E^e]$ matrices (used in the calibration as well as 'Case 1' of the device measurement) has the effect of greatly improving measurement accuracy. The lengths corresponding to a maximum determinant over the range of measurement frequencies can be easily and reliably found using standard optimisation techniques. Figure 4.6 shows simulated de-embedding of a particular device parameter, $S_{BA}^{DUT}(2,2)$, in a multimode device. In this simulation, the matched loads and probes used in the calibration and device measurements are non-ideal and the system is disturbed by random measurement noise with an amplitude of up to 1dB. Clearly, the accuracy is much higher in the system with higher values of $|E|$. Note that in this case the system was determined with the number of lines in both cases exactly equal to the number of modes, i.e. not over-determined.

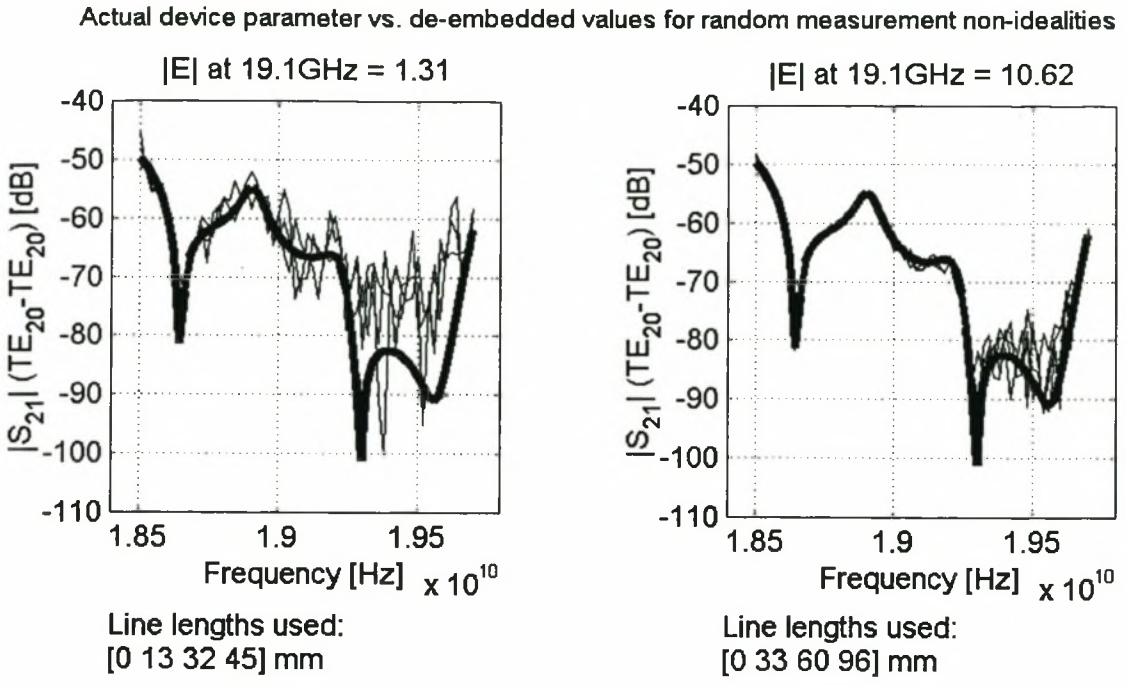


Figure 4.6. Simulated de-embedding of a device parameter with different choices of line lengths under non-ideal measurement circumstances.

4.5.2 Effects of non-ideal loads and probes

To investigate the effects of non-ideal probe sections and non-ideal matched loads, these parameters were perturbed randomly during simulation, after which the probe and device parameters were extracted from the resulting simulated measurements, according to the methods presented here. Figure 4.7 shows the gradual decline in measurement accuracy of two example device parameters as the probe and load reflections are increased. Note that the chosen device parameters include very low values of $\left[S_{BA}^{DUT}\right]$ which are extremely difficult to measure. These simulations thus represent a stringent test of the method's capabilities.

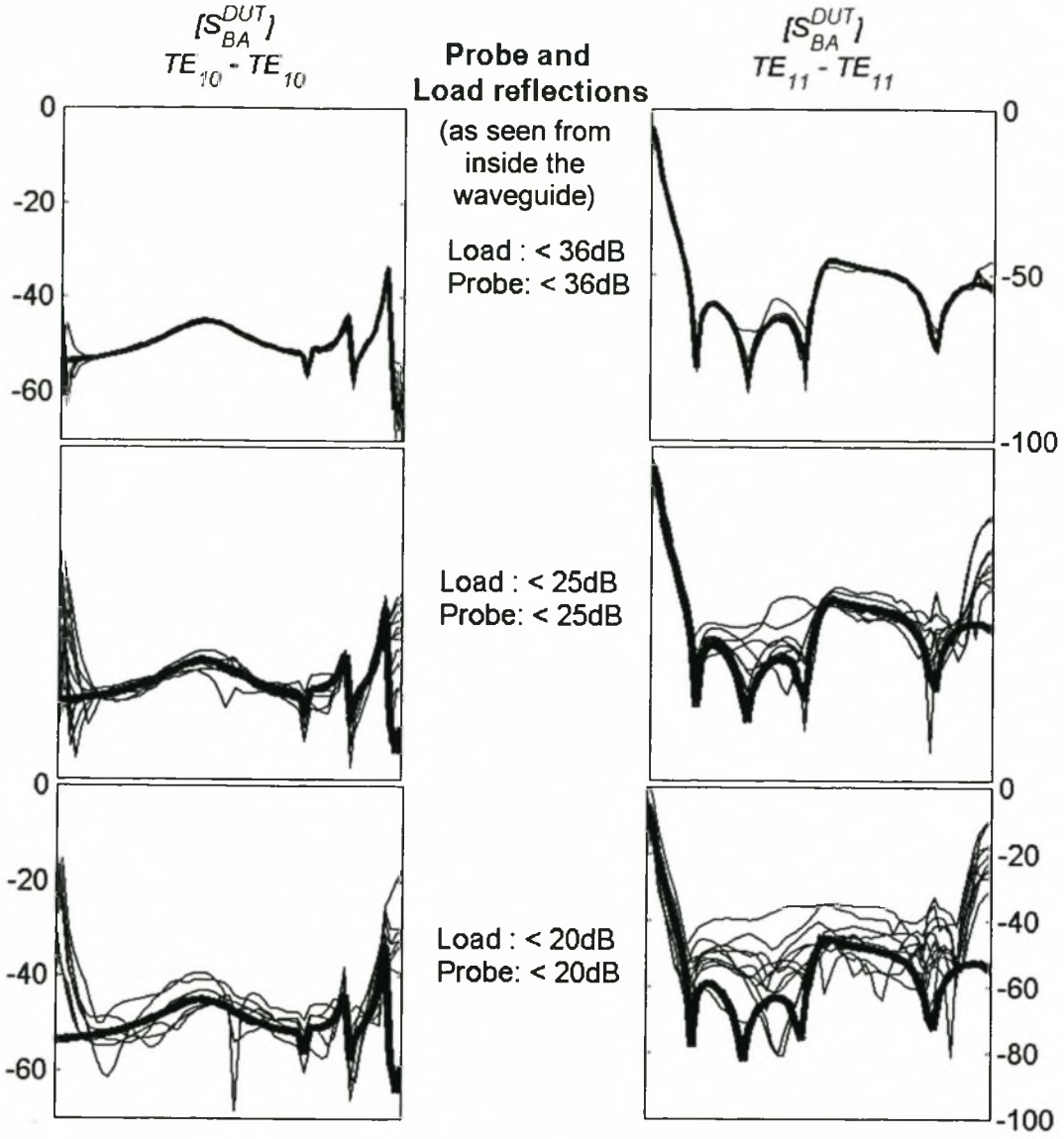


Figure 4.7. Effect of random simulated load and probe non-idealities on sample device parameters

4.5.3 Effects of using an over-determined system

As the construction of an over-determined solution involves more measurement, it is important to investigate its improvement in accuracy over a standard determined system. For this test, the second device measurement technique was used to de-embed parameters from simulated measurements, for a number of random devices and transitions. The iterative solver was given a

number of random initial points with a maximum absolute S-parameter error of about 0.4. Figure 4.8 shows the convergence of the iterative solver to reduce the maximum absolute error (over all device s-parameters). Clearly, the use of more measurements, or line combinations, results in faster and more accurate performance.

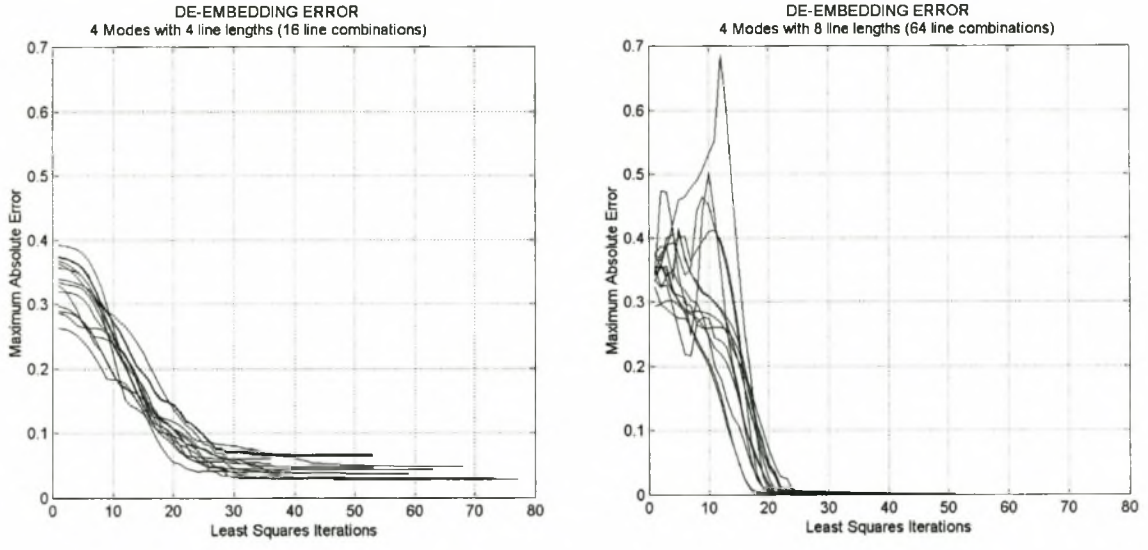


Figure 4.8. Simulated device de-embedding for random device and transition parameters. Each line corresponds to a different starting value set.

4.6 Measurement of a Multimode Bandstop Filter

For the purposes of verifying the measurement techniques, a multimode bandstop filter was constructed in WR-90 at 19GHz to attenuate all incident modes by at least 30dB over a 4% bandwidth. These filters find application as chokes in microwave heating applications. Their design is detailed in chapter 5. Five modes are allowed to propagate at this frequency, two of which are TE_{11} and TM_{11}, the only modes that are cross-coupled by the device. This device falls into the 'special case' category discussed in the previous sections, and the corresponding simplified techniques were used. As such, these measurements do not verify the most general case. This verification was, however, done by computer simulation and the results are shown in figure 4.8.

Analysis of the choke devices is done with a combination of the Mode Matching technique and generalized scattering matrix, as discussed in chapter 3, and is presented in figure 4.9, together with the values extracted from the measurements. As the transmission parameters are the only ones of importance in this case, only they are shown. The device, also shown in figure 4.9, was milled to an accuracy of 50μm from aluminium.

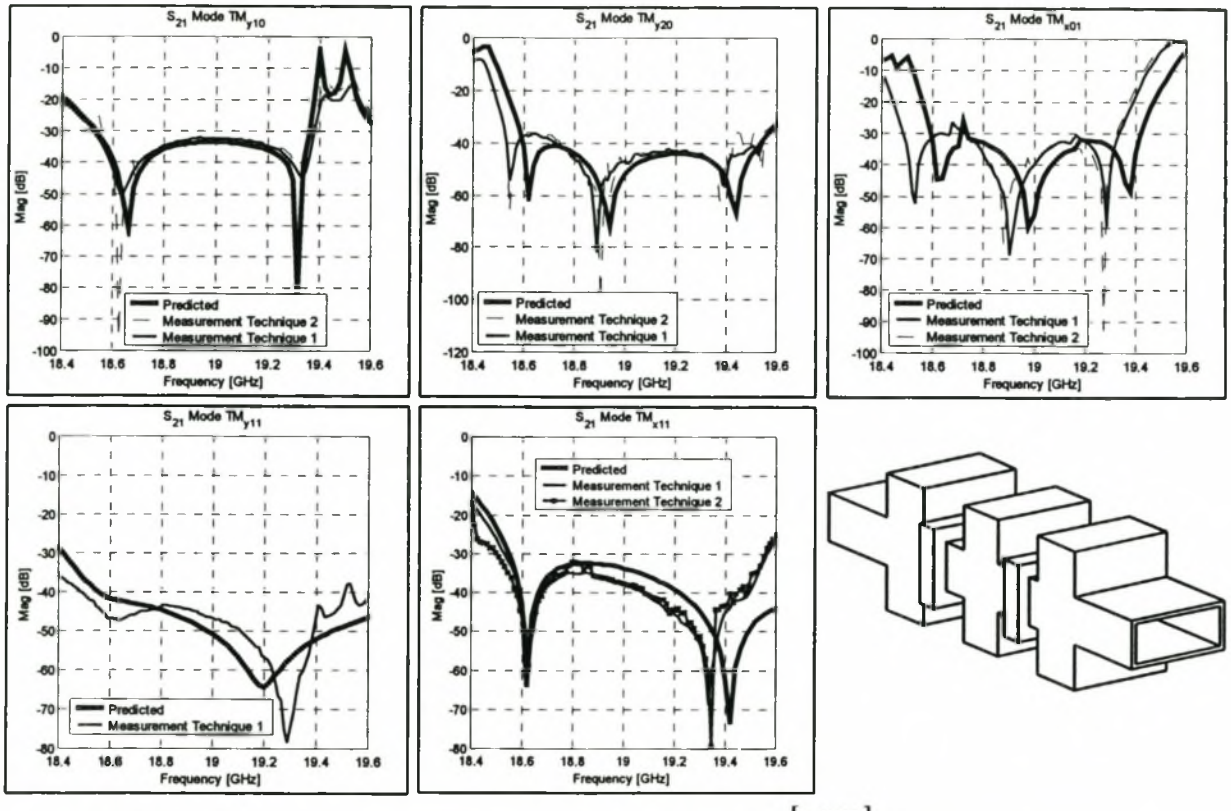


Figure 4.9. Predicted and Measured diagonal elements of the $\left[S_{BA}^{DUT}\right]$ matrix with both techniques and a drawing of the device that was measured

Measurements were done on an HP8510 VNA with three different probe positions to extract the parameters for all five modes. The predicted device parameters were converted from TE_z/TM_z format to TM_y and TM_x using the techniques described in chapter 3 for the purposes of comparison. The results in figure 4.9 show clear agreement with the predicted parameters and between the two different techniques. Slight frequency shifts are due to manufacturing limitations, which at 50μm would cause such 0.5% deviations. 'Measurement technique 1' uses two load-

terminated transitions, and 'Measurement technique 2' uses one load terminated transition and one sliding offset short transition.

4.7 Conclusion

A scheme for multimode device measurements in overmoded waveguide has been developed. The technique uses variable line lengths in different configurations with lightly coupled waveguide probes to exploit the differences in propagation constants between the modes. Degenerate modes are accounted for by working with TE_x/TM_z or TE_y/TM_y modes sets with applicable probe positioning.

The full transition calibration procedure is detailed and two approaches to device measurement are discussed. The first approach results in a simple linear system of equations to solve for the device, while the second requires the solution of a larger system of non-linear equations, but provides considerable measurement advantages.

Both methods are used to measure a sample multimode device and the results are shown to agree well with theoretical predictions, to within a few dB at levels of -50dB and lower.

Chapter 5 - Design Solutions: Overmoded Band Stop Filters

5.1 Introduction

Design algorithms for single-mode waveguide filters have been in existence for decades and can be regarded as a mature field of study. Recently, however, developments in commercial microwave applications, such as dielectric heating, have posed new problems, specifically in the design of filters in waveguides supporting multiple propagating modes. To date, overmoded filter design has enjoyed very little attention from researchers, primarily because of the complex nature of the topic and the strong divergence it exhibits from conventional network synthesis-based filter design techniques. In 1968, Butterweck proposed a series of 'mode filters', incorporating suitably positioned resistive sheets in oversize waveguide intended to impede the propagation of all but the TE_{10} or (more generally) TE_{m0} modes [Butterweck 1968]. Later, Ren et al. presented a design technique for overmoded circular waveguide filters that uses the image method to position variable-length, axially-oriented disks of dielectric in the waveguide [Ren 1974]. A limitation of the design, however, is that it relies on the modes having the same propagation speeds, which is seldom possible, especially in rectangular waveguide. Research further afield, on overmoded propagation (discussed in chapter 1) focuses primarily on mode converter and horn designs, and sheds little light on potential approaches to filter design in the presence of multiple propagating modes. Multiple modes are, however, widely exploited in conventional filter design, [Atia 1971] where a cylindrical cavity may be designed to support two or more modes at resonance. This effectively provides additional resonators for 'free', resulting in smaller and lighter filter topologies. The use of oversize waveguide in this way also provides higher Q values and associated lower loss. The design philosophy behind such filters is, however, essentially single mode, and the exploitation of multiple propagating modes is as a result of a choice of topology and limited to very narrow band applications.

In contrast, this chapter concerns itself with devices with physical port sizes that support more than one propagating mode and must exhibit filtering properties for more than one

mode at once. In such cases, transmission through the filter is described by a $[S_{21}]$ matrix, defining each mode's propagation through the filter, as well as the couplings between modes that might occur due to the filter's structural characteristics. If the device supports M modes at its input port, and N modes at its output port, it can be described by modelling each mode as an equivalent electrical port, resulting in an $M+N$-port structure. Filter specifications must then be made with this $M+N$-port structure in mind. As discussed in chapter 1, this type of structure clearly exhibits properties that make it difficult to design using conventional network synthesis and lumped element modelling-based methods.

This chapter introduces a novel, systematic method of designing stop band filters in the multi-mode environment based on the CAD principle of Structural Growth with Functional Blocks, outlined in chapter 2. There are no restrictions on input/output-port waveguide sizes and geometries or on maximum attenuation and bandwidth, though the size of the filter may increase with more stringent specifications. The design procedure allows the specification of individual stop-bands for the various modes and minimizes cross-coupling through the choice of suitable functional block components. The developments presented here are an original contribution to the field [Vale 2000b, Vale 2000c] and are inspired by the need for an effective solution to the problem of reactive choke design in the field of industrial microwave heating [Vale 2000d]. This chapter will start by discussing this industrial application, which imposes certain limitations on the choice of functional blocks and the filter's general topology. This topology, as well as the type of functional blocks used will be discussed in the next section. Next, the operation and selection of the functional blocks will be detailed. This will be followed by the filter design using the strategies discussed in chapter 2. Finally a number of examples of measured and designed filters will be present.

5.2 An Industrial Application: Microwave Heating

In many microwave heating facilities, conveyor belts passing through a microwave heating cavity enforce permanent large openings in the cavity sidewalls, through which dangerous levels of microwave energy can escape if not filtered. The solution to this problem is the placement of 'chokes' on either side of the cavity to reflect and/or absorb any energy before it can do harm to people or equipment outside. These chokes must at the same time allow the free movement of product through the facility.

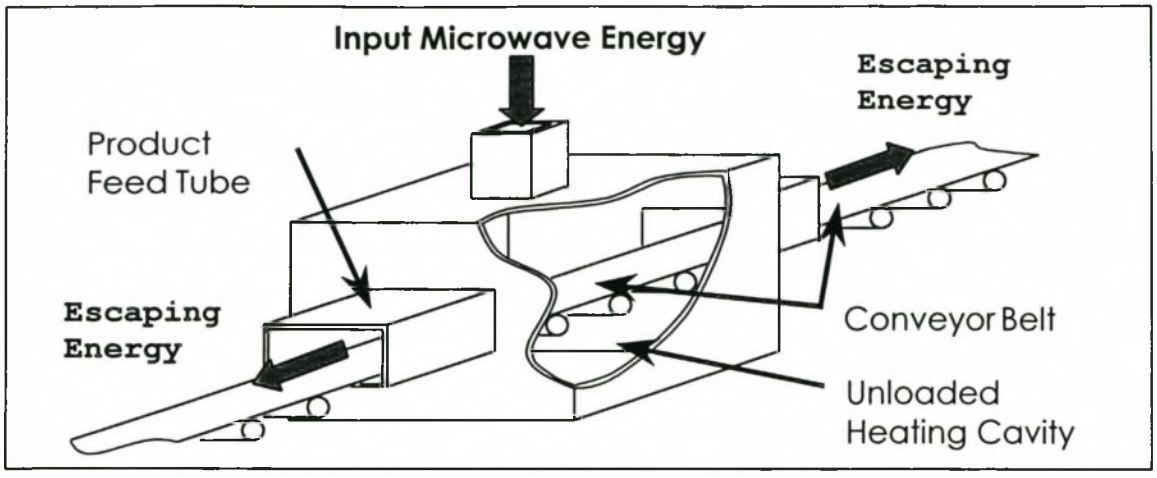


Figure 5.1 A microwave heating cavity with conveyor belt feed

The preferred choke design solution is a reactive choke. This typically reflects the escaping energy back into the cavity using equivalent reactive elements, such as stub lines, in a similar fashion as would be used in a conventional bandstop filter design. Because of the complex nature of the microwave design problem, however, standard reactive choke designs typically enforce limitations on the aperture geometry, often requiring it to be rather small, or else wide but very thin [CMPM 1994, Soto 2000] so as to apply monomode equivalent approaches. When such limited aperture geometries conflict with the physical requirements for aperture size set by the size of the product feed tube, the designer must resort to other solutions such as tunnels with absorbing walls, special arrangements of doors timed to open and close to admit product or 'maze openings' that force the product to 'meander through a folder corridor lined with absorbing walls' [CMPM 1994]. These approaches have many drawbacks, specifically since the use of absorbing materials requires bulky cooling and imposes power limitations, and the free flow of product may be impeded by obstacles such as doors and chain walls. Alternative empirical designs have the drawback of being limited to specific geometries and require redesign from scratch for each new situation.

A design procedure for overmoded bandstop filters effectively provides a more flexible solution to the design of reactive chokes with arbitrary aperture geometries, provided the functional blocks used do not intrude into the aperture. The choice of functional block that fulfils this requirement, as well as the proposed topology of the filter are discussed in the next section.

5.3 Basic Filter Structure

Beyond considerations made for the intended application of the devices, a major factor governing the shape of the filters is cross-coupling between modes at functional blocks. Designing a bandstop filter under such circumstances is highly problematic, as a functional block which creates a null for one mode will excite a number of other modes at the same frequency. The mode matching equivalent circuit derived in chapter 3 indicates that cross-coupling between modes of different propagation constants could be greatly reduced by using symmetrical functional blocks consisting of back-to-back waveguide steps separated by lengths of uniform guide. If the steps go from small to large to small again, and the size of the small waveguide is the same as the product feed tubes discussed above, then such functional blocks in cascade would fulfil the requirements of a reactive choke structure in terms of geometric limitations. Investigation has found that such functional blocks can, with the correct choices of step size and separating length, be made to provide a propagating null for one or more modes. Their principle of operation will be discussed in the next section.

The general filter topology can now be projected as shown in figure 5.2.

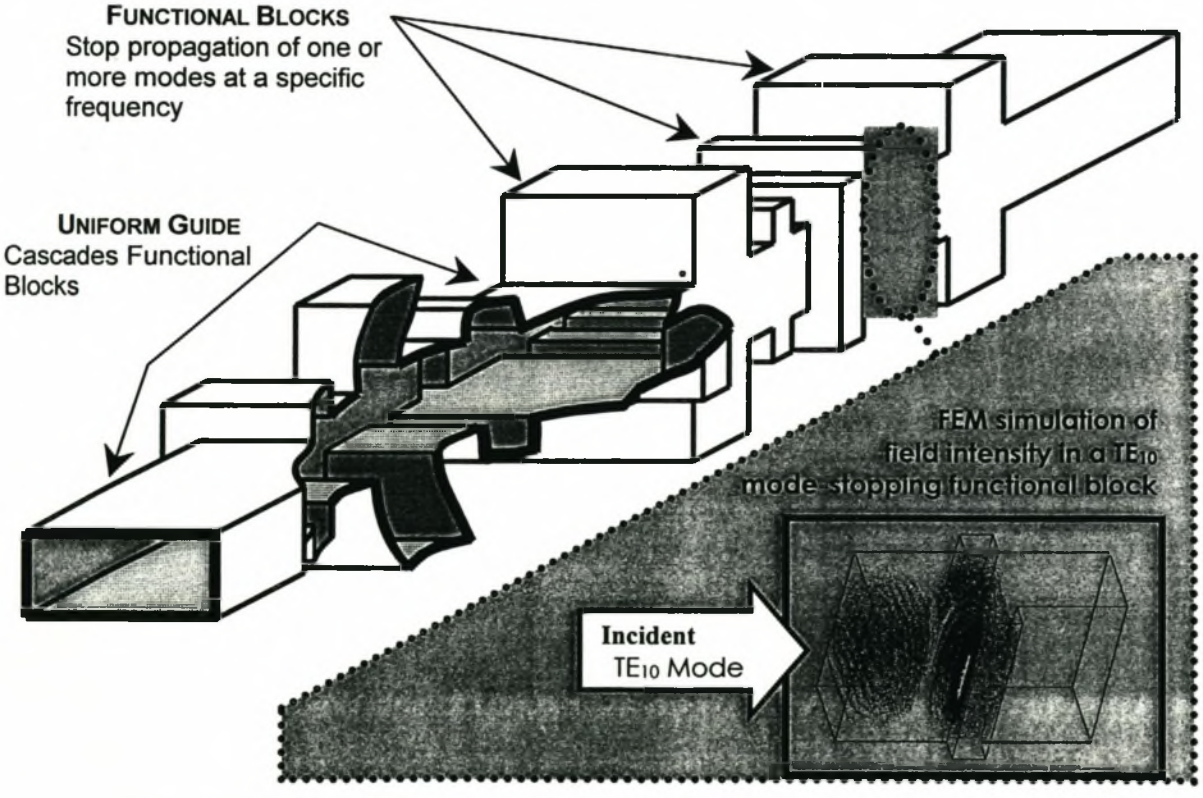


Figure 5.2 Proposed Overmoded Bandstop Filter/Multimode Choke Structure

This structure relies on the fact that a stop band can be built up of a number of separate functional blocks in cascade. This operation is shown for the TE_{10} mode in figure 5.3. The distance separating the blocks is critical to the stop band response and is a design parameter that will be discussed in more detail in section 5.5.

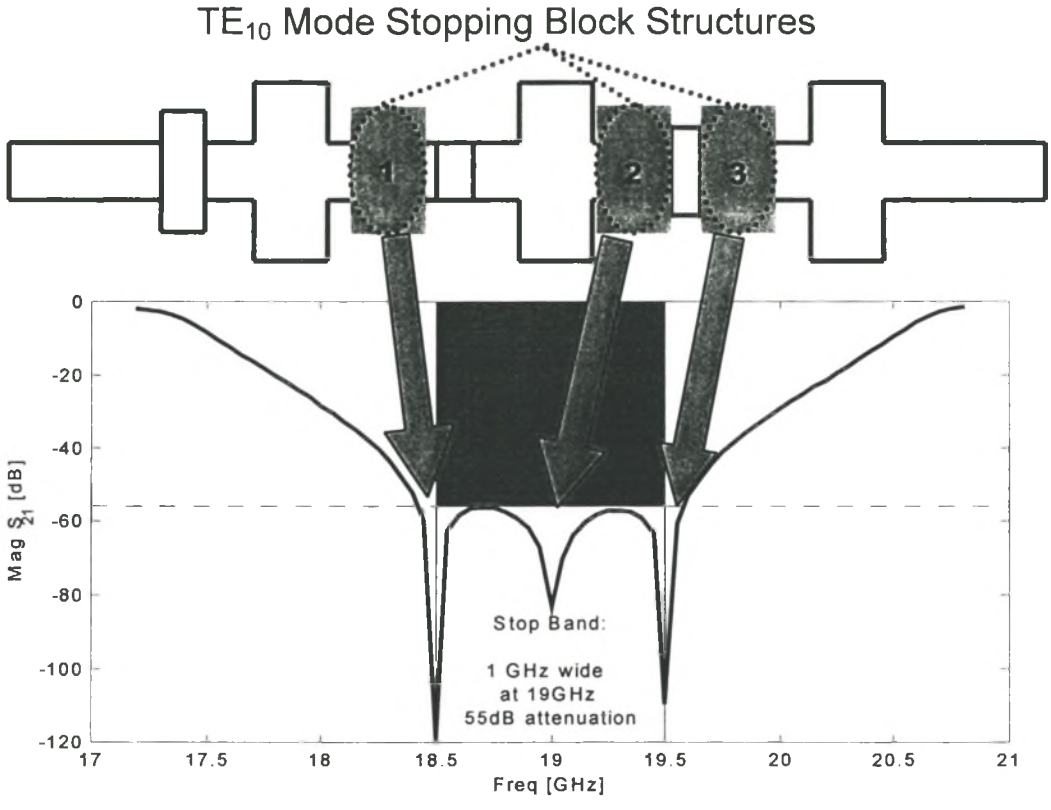


Figure 5.3 TE_{10} Mode attenuated by a system of three Functional Blocks

In order to move from a proposed design structure to a working prototype, there are a number of requirements of both the functional blocks and of their cascaded arrangement. There are also some short cuts that can be taken by allowing some functional blocks to operate on more than one mode at a time.

5.4 Selection of the Functional Blocks

As discussed in chapter 2, the first stage of the growth-based CAD strategy is the selection of one or more functional blocks to make up the structure. As has been discussed above, the emerging functional block consists of back-to-back small-to-large double- or single-plane steps in waveguide, where the small waveguide dimension is the same as the input/output

port aperture size. The block can therefore be defined by three design variables, width (a), height (b) and length (l) as shown in figure 5.4.

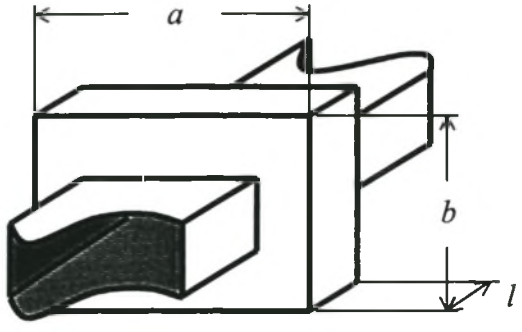


Figure 5.4 Functional Block – Mode-Stopping Discontinuity

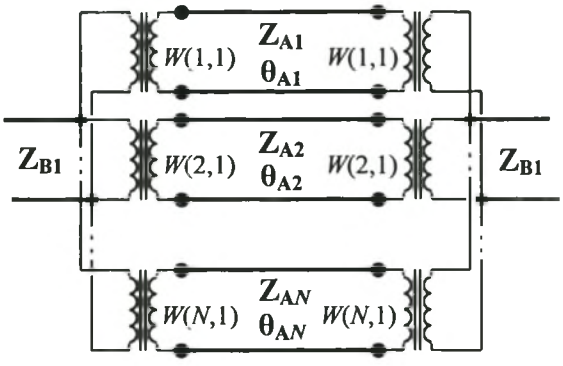


Figure 5.5 Approximate circuit of functional block

The fundamental field activity in this block can easy be understood with the aid of the mode matching equivalent circuit. Consider a mode matching operation done on the small to large step comprising the functional block with N modes on the large side, and M modes on the small side. Now if the small side modes are truncated to M=1, an approximated circuit can be drawn for the functional block as in figure 5.5. In geometries which exhibit resonance, incident energy in the small guide is found to be split relatively evenly between two modes of different propagation constant in the large waveguide. The electrical length of the large guide is different for the two modes, and, if chosen correctly, can allow destructive recombination of the energy in the two modes at the opposing step, leading to a null in propagation at a specific frequency. Since the dimensions of the waveguide step determine both how the energy is split between the two carriers and what their propagation constants are, and the length of the enlarged guide sets the path length that the modes must travel before recombining, it is clear that the performance of the block depends on a rather complex relationship between dimensions.

It would be quite possible to use this general functional block definition, with a, b and l as design variables, and apply a growth-based CAD strategy as outlined in chapter 2, to determine both the design variables of the multiple instances of these blocks and their arrangements and separations. This approach would, however, be unnecessarily inefficient and potentially unreliable given the prior knowledge available.

It is known that the functional blocks must be tuned to resonance, where they provide propagation nulls for various modes. It is also known that these nulls should occur within the stop band of the filter. By investigating each mode in isolation, a first order approximation can be made for the number and position of the nulls based on the specified bandwidth and in-band attenuation. Lastly, an approximation can be made that cascading of functional blocks will not alter their resonant frequencies. This approximation relies on sufficiently large separation distances between adjacent functional blocks so as to ensure non-interaction of localized modes.

It would be a waste of time, if multiple instances of resonators of a very similar nature required individual tuning from scratch by the CAD procedure to resonances that are already known by the designer. The large amount of prior knowledge about the functional block resonances can be exploited to overcome this inefficiency. This is accomplished by defining a number of functional blocks as above, each with a, b and l predefined by the designer. In terms of the CAD strategy used, the functional blocks will therefore have no design variables and need only be cascaded appropriately. This approach has the additional advantage that the analysis of the functional blocks is trivial, as they are fixed and need only be analysed once: upon definition. In this particular case, then, the functional block selection stage is somewhat more involved than usual. The rest of this section details two methods that can be used to assemble the necessary suite of predefined functional blocks. The first relies on a simple grid search technique, and the second exploits the available knowledge of how the blocks work in terms of the approximate circuit in figure 5.5.

a) A Grid Search Method

Since the functional block is described by only three variables, it is quite possible to use a constrained grid search method. It should be noted that large cross-sectional dimensions often result in multiple resonances in the band, which could make for a shorter filter, but will complicate the cascading process. The search typically results in clouds of viable geometries for each mode. An example for the TE_{10} mode in WR-90 at 19.1GHz is shown in figure 5.6.

Ideally, the block would be chosen to be as short as possible, though it is clear that a high sensitivity to the height dimension results from such a choice. At the cost of increased block length, a better choice would be a geometry from higher up on the map, where the gradient in terms of the b dimension flattens out somewhat.

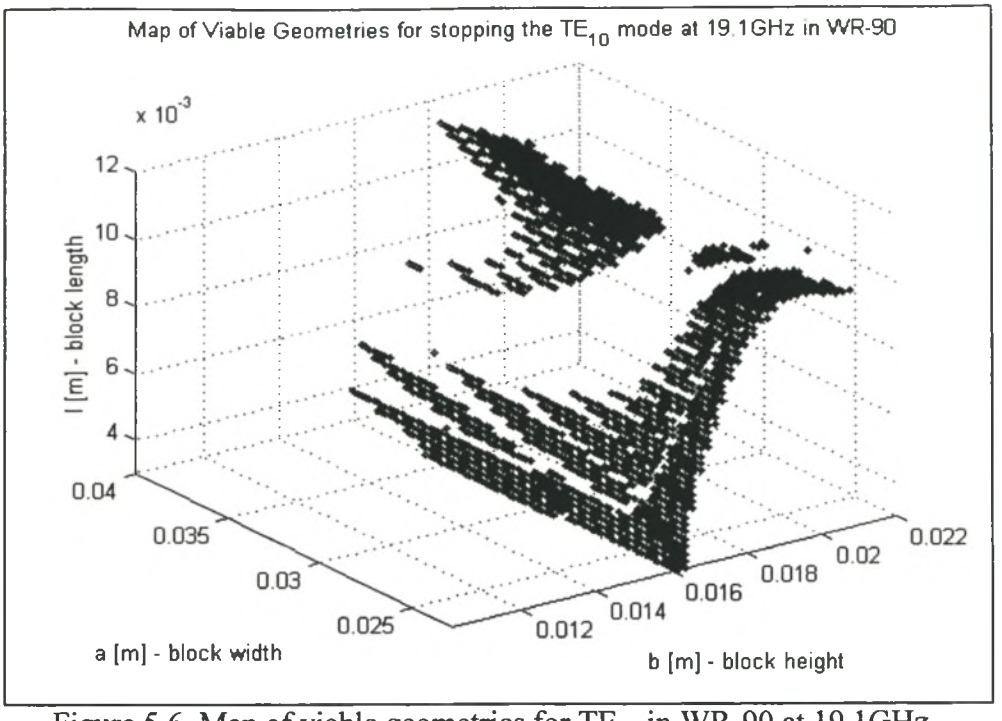


Figure 5.6. Map of viable geometries for TE_{10} in WR-90 at 19.1GHz

Also, a general petering out of viable geometries as the *a*-dimension increases, indicates areas where the propagation zero might not be well defined, or the roll off might be too sharp or jagged, and thus not particularly suited for the bandstop filter. More useful geometries can often be found in denser areas.

One other important use of the map is to identify blocks that can be used for more than one mode simultaneously, reducing the number of blocks and the size of filter greatly. This can be accomplished by superimposing two maps over each other and selecting geometries where the two maps overlap or come close to each other. All the designs presented here use some blocks which exhibit this property. This practice complicates the cascading process somewhat, as two such blocks can seldom be separated by a distance that favours all the resonant modes by reducing propagating between the nulls as is desired.

b) The [*W*] Matrix Method

The principle of destructive signal recombination described earlier can be used to derive a more elegant alternative method of designing the functional blocks. This method is based on an observation that the incident energy from the primary waveguide should split roughly evenly between only two modes in the block for optimum cancellation. The higher of these two modes can be either propagating or evanescent.

To find geometries that accomplish such a split, the ratio of coupling, R_c, (chapter 3, eqn. (3.36)) can be investigated at the filter's centre frequency between the mode of interest in the small guide and the various modes in the large guide. Assuming that it is desired to stop mode m, a number of values $R_c(i,m)$, $i = 1..N$, are computed for a number of modes, i, in the larger guide for a sweep of the variables, a and b. $R_c(m,m)$ is typically large, and the aim is to find a cross-section (a,b) for which all other values are small, except for one other mode, j. A figure of unsuitability for resonance for various geometries is shown in eqn. 5.1,

$$I_{mj}(a,b) = \left|\alpha_m R_c^{(a,b)}(m,m) - \alpha_j R_c^{(a,b)}(j,m)\right| + \sum_{i=1, i\notin(m,j)}^{N} \alpha_i R_c^{(a,b)}(i,m) \tag{5.1}$$

Here $R_c^{(a,b)}$ is the ratio of coupling for step geometry, (a,b) and N is the number of modes considered in the large guide. α_k is a constant determined by the attenuation of mode k in the large guide at the centre frequency over a length of the order of the smaller guide's dimensions (The exact length is not crucial, as long as it is within a factor 10 of the eventual length). Its purpose is to eliminate the effect of non-propagating modes that would not have a strong influence over the length of the guide. Note that modes with the same propagation constant, $j=k$ and $j=k'$, must be grouped together by adding their ratios of couplings. For a sweep of values a and b, I can be computed very quickly. Lower values of I_{mj} indicate more suitable geometries for resonance due to the interaction of modes m and j.

Figure 5.7(a) shows a landscape plot of I for creating a null in mode TE_{20} with the aid of modes TE_{22} and TM_{22}. The higher points correspond to geometries that will not lead to a resonance for that mode at any frequency, and the lowest points correspond to the most probable resonant geometries. Once a step geometry is determined, the length of the block can be calculated by iteration from a starting value that causes modes m and j to fall 180 degrees out of phase. To illustrate how the two methods lead to the same viable values (a,b), the map produced by the grid search is shown with the l-dimension removed, in figure 5.7(b). Close inspection of figure 5.7(a) shows that resonances reported by the grid search method concur with values of I below some critical value, normally around 0.7. Note that $[W]$ matrix method works over a rather large band of frequencies, implying that block structures for a particular mode will generally be of the same shape. Choosing geometries that correspond to low values in figure 5.7(a) will therefore lead to good resonant blocks.

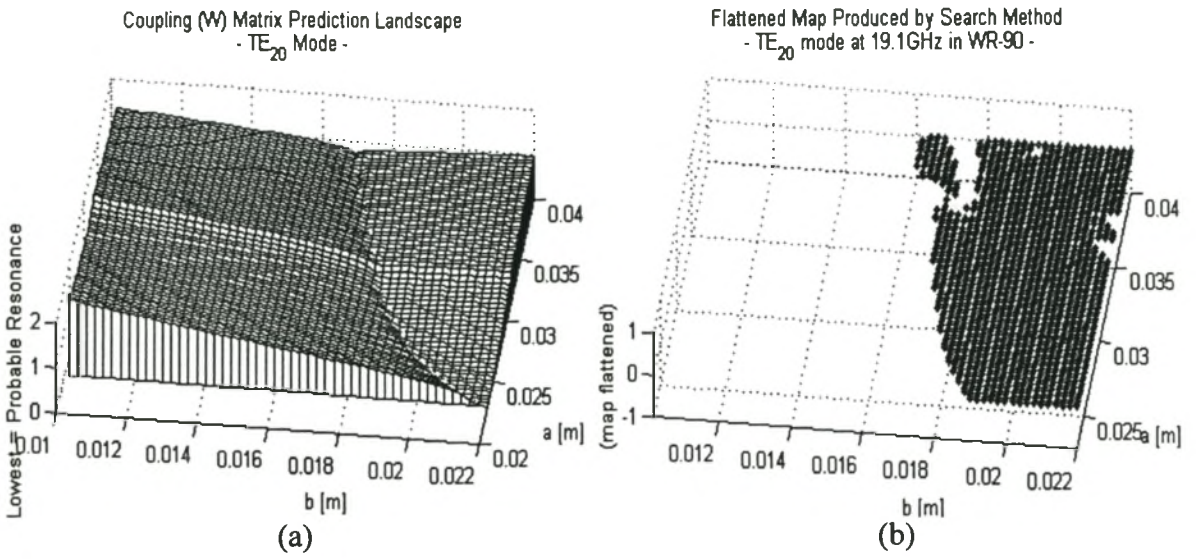


Figure 5.7. Comparison of resonant geometries predicted by [W] matrix and the grid search methods

Because computation of the [W] matrix involves no matrix inversion or frequency dependence, this method is much quicker than the grid search method.

Whatever method is used, once a suitable geometry is found for a functional block, an iterative procedure can be used to tune that geometry to the correct resonant frequency for one or more of the propagating modes. Copies of similar blocks can emerge, each tuned to slightly different resonant frequencies. The result of this phase of design should be a series of functional blocks, each responsible for a propagation null of one or more modes at some point in the stop band. All the nulls provided by all the blocks should be enough to build up a suitable stop band for all the individual modes. This may require a relatively large number of blocks, depending on the specifications.

5.5 Filter Design using Growth-Based CAD Strategies

In order to motivate the use of growth-based CAD strategies for this design problem, consider the requirements for effectively cascading the functional blocks comprising the filter. There are two criteria that influence the optimal choice of waveguide length between adjacent functional blocks.

The first is that blocks that resonate for the same mode must be separated by an electrical length that ensures maximum attenuation at the frequencies between the two resonances. If zero cross-coupling is assumed, the transmission of mode m through two cascaded functional blocks, A and B, can be written as in eqn. 5.2, where β is the propagation constant of mode m in the uniform guide separating the blocks, and l_{AB} is the separating length. S_{xy}^{Z} refer to the GSMs of the functional blocks. To maximize attenuation as is desired, design condition 1, in eqn. (5.3), must be satisfied.

$$s_{21} = \frac{S_{21}^{A}(m,m)\cdot S_{21}^{B}(m,m)\cdot e^{j\theta}}{1 - S_{22}^{A}(m,m)\cdot S_{11}^{B}(m,m)\cdot e^{j2\theta}} \qquad \theta = -\beta \cdot l_{AB} \tag{5.2}$$

$$\text{Condition 1: } \angle\left[S_{22}^{A}(m,m)S_{11}^{B}(m,m)e^{j2\theta}\right] = \pi \tag{5.3}$$

If another block, C, separates A and B, and does not resonate for mode m, it can be accounted for as an effective phase shift by adding the angle of $S_{21}^{C}(m,m)$ to θ in eqn. (5.2) above.

The second criteria enforces a minimum separation length between adjacent blocks to ensure that localized modes do not interact and disturb the blocks' resonances. This measure is necessitated by the approximation made in the functional block selection stage, where it is assumed that functional blocks can be independently tuned to resonance before being cascaded. This condition is expressed in eqn. (5.4) where α_i is the attenuation of any non-propagating mode excited by block A or B.

$$\text{Condition 2: } \exp(-\alpha_i l_{AB}) << 1 \tag{5.4}$$

If each functional block was selected so as to resonate for only one mode at one frequency, then it would be possible to use the criteria above to directly cascade the functional blocks. In practice, however, the overall length of the filter must be kept to a minimum. Simply applying the criteria above often leads to long lengths of uniform guide between functional blocks resonant for one mode that could be filled with a functional block resonant for another mode, saving considerable space. Furthermore, the practice of using functional blocks that resonate for more than one mode provides reductions in complexity and size of the filter. This

does, however, come at a cost. Inevitably, such arrangements lead to conflicting choices for optimal separating lengths between adjacent blocks. Phase compensation can be introduced by another block inserted between two such conflicting blocks to solve the problem. Finally, an added complication arises when cross coupling *does* occur between propagating modes. This causes eqn. (5.2) to become invalid, and makes the optimal cascading length very difficult to find analytically.

Attempting direct synthesis of the filter under such circumstances can become tedious and often leads to unnecessarily long and inefficient designs. This is because of the large number of possible arrangements of functional blocks and the degree to which cascading length choices often rely on compromising between conflicting values. These complications prompt the use of a growth-based CAD strategy to find the most appropriate arrangement of functional blocks and separation lengths. Two different strategies are found to work effectively on this design problem. They are the simplified tree search strategy and the evolutionary strategy.

a) The Tree-Search Strategy

This strategy is particularly suited to this problem, due to the availability of the cascading criteria listed above, which allows the insertion of the functional blocks at the end of the structure without any optimization through the use of prior knowledge of the problem. In addition, since it is known that each functional block need only be used once in the design, the extent of tree search strategy can be limited by constraining each block to be only used once in the structure, according to the suggestions in chapter 2. Under such circumstances, the simplified tree search strategy becomes much more efficient. Since its application here differs somewhat from the general algorithm proposed in chapter 2, it is worthy of a more thorough discription. A diagram of the search as applied to a 4-block problem, is shown in figure 5.8.

The algorithm uses one of the available functional blocks as a starting structure. It then attempts to cascade each of the remaining blocks to this starting structure. The principle behind the search is that only a selection of these blocks will successfully cascade, and the rest will be rejected. In this problem, a block's viability is easily assessed by considering the conditions in eqns (5.3) and (5.4). Only those blocks for which these conditions predict non-conflicting cascading lengths are allowed to form a new branch of the search.

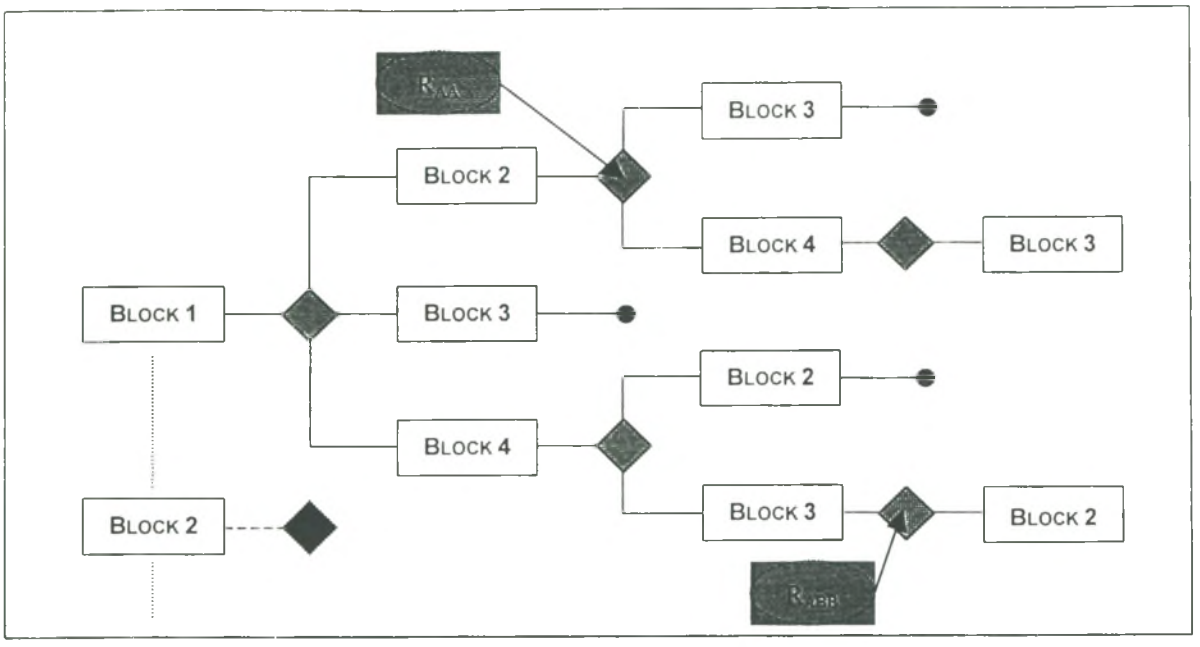


Figure 5.8. Schematic of the tree search algorithm

To favour short filter structures, the search can be further restricted by disallowing excessive cascading lengths. Each time a block is successfully placed, a new branch is formed and that block is removed from the pool of available blocks for that branch. Should it occur at some point that none of the available blocks fits, that branch of the search is 'killed' off (—●) and the algorithm steps back a level to attempt the remaining untried branches. Should all the blocks be successfully placed in a branch, then a viable solution has been found and can be stored. Typically a number of viable solutions are found, depending on the stringency of the cascading requirements. Stricter cascading rules lead to fewer solutions, but typically of a higher quality.

The advantage of using the strategy in this way is that analysis of the structures is unnecessary at all points in the search except at the end. This makes the design extremely fast, though sometimes unreliable, as it is based solely on the cascading conditions. An alternative implementation retains the structures' responses as they are built up, as described in chapter 2 for this strategy. Rejection of blocks is then decided on both by the cascading conditions and by the influence the proposed block would have on the structural response. This leads to a much more reliable search at the cost of a more computationally expensive procedure.

When compared to a manual trial-and-error based design, this procedure is not only many times faster, but also produces significantly better results. Furthermore, since the search can be instructed to stop searching a branch if the structure already fulfils the specifications of the design, it is possible to provide redundant functional blocks and allow the procedure to select the most suitable among their number. In many cases, solutions produced by this method utilized the functional blocks more effectively than the by-hand designs, allowing fewer blocks to be used and making for much shorter filter structures.

b) The Evolutionary Strategy

As a growth-based CAD procedure, the evolutionary strategy is tremendously flexible. As discussed in chapter 2, this procedure effectively uses evolutionary principles to grow microwave structures from single functional block 'organisms'. The genetic equivalent of 'reproduction' is accomplished by cutting and splicing parent structures and/or simple addition or subtraction of a functional block or changing of a connecting length. 'Competition' is accomplished by pairing random structures and, in this problem, eliminating the contender whose peak in-band transmission over all the modes was the highest. To encourage shorter structures, contenders with similar performance were judged instead on the basis of overall length.

In contrast to the previous method, this approach does not take into account the cascading conditions listed earlier, and requires full analysis of the growing structures as the process continues. The result is a considerably slower design procedure, but with massive benefits in terms of flexibility and robustness. This is mainly because the technique does not make any assumptions about cross coupling and the interaction of localized modes – rather it takes into account such features through rigorous analysis, and consequently allows them to potentially aid the goals of the design. An added advantage is the ability to include simple phase compensation blocks, that can be naturally selected along with the other functional blocks to provide improvements in performance.

The algorithm was successfully used on a variety of designs, some of which are presented in the next section. Details of the design of a scaled version of a 2.45 GHz industrial system at 19.1 GHz filter are shown here by way of example. Figure 5.8 shows a plot of the peak $|S_{21}|$ over all frequencies and propagating modes of the best member of the population for each generation. Considering that the lowest possible value is –41dB, the fact that the genetic

algorithm reaches this value to within only a few dB's before levelling off confirms its appropriateness for this problem. Figure 5.9 is a histogram of the peak $|S_{21}|$ distribution over the best 500 members of the 1000 strong populace after 100 generations. Clearly, most of the filters/populace have evolved to yield a peak value of close to -40dB at this time. Figure 5.10 is a histogram of total filter length distribution after 100 generations. Most filters have settled to a length around 110mm, which is the same length as that obtained using the non-evolutionary method discussed above.

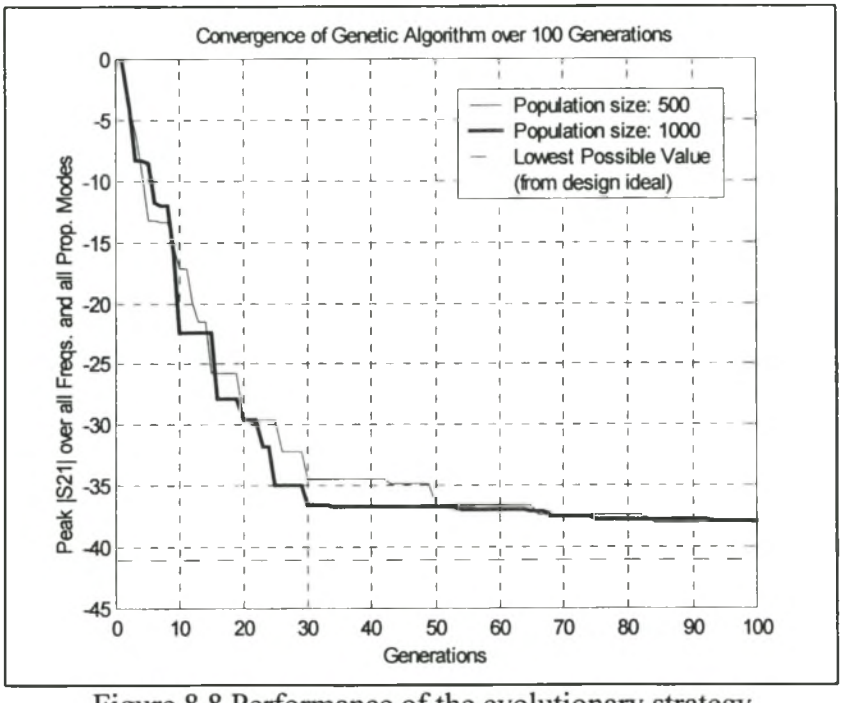


Figure 8.8 Performance of the evolutionary strategy

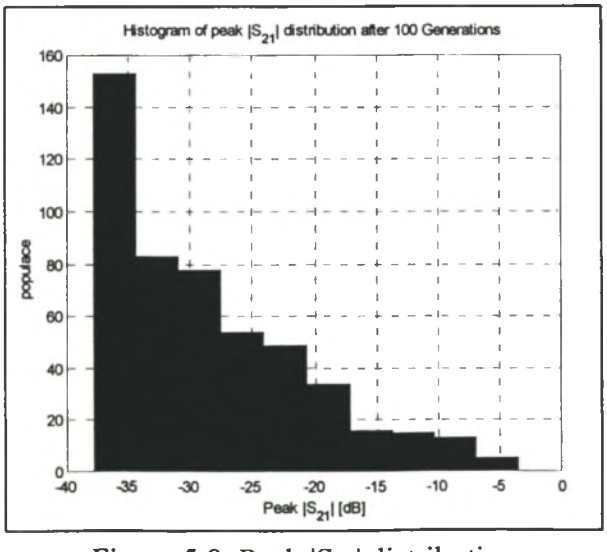


Figure 5.9. Peak $|S_{21}|$ distribution

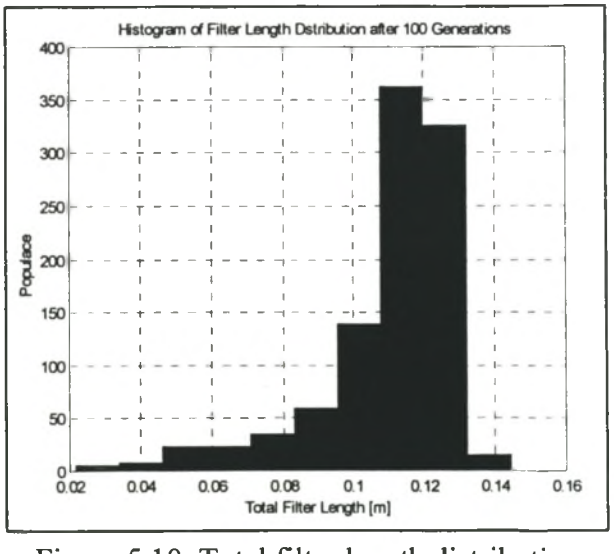


Figure 5.10. Total filter length distribution

It is interesting to note that the tree-search and evolutionary strategy frequently produced very similar design solutions, suggesting that both techniques may have been approaching near optimal solutions considering their diverse design philosophies.

In both cases, standard structural optimisation was always used on the solutions produced by the design procedures. This was typically to refine them to specification or to adjust certain dimensions slightly (for the purposes of manufacture). The time spent on this was typically rather small, as the designs were usually good.

5.6 Design Examples

FILTER 1. (19.1GHZ BANDSTOP FILTER)

This filter was designed for verification of the design and analysis techniques by physical measurement. The centre frequency is 19.1GHz and the bandwidth is 5%. The filter is designed in WR-90, with input and output port sizes of 22.86mm by 10.16mm supporting five propagating modes throughout the band. The filter is to provide at least 40dB attenuation throughout the band for all five modes

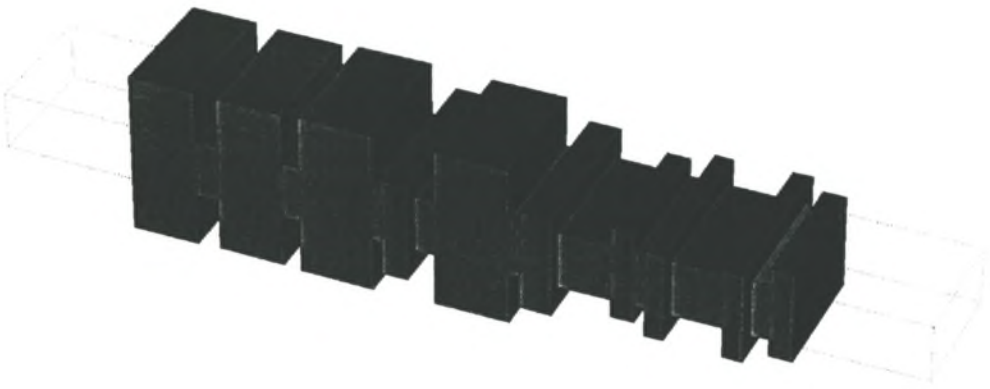


Figure 5.11 Schematic of Filter 1

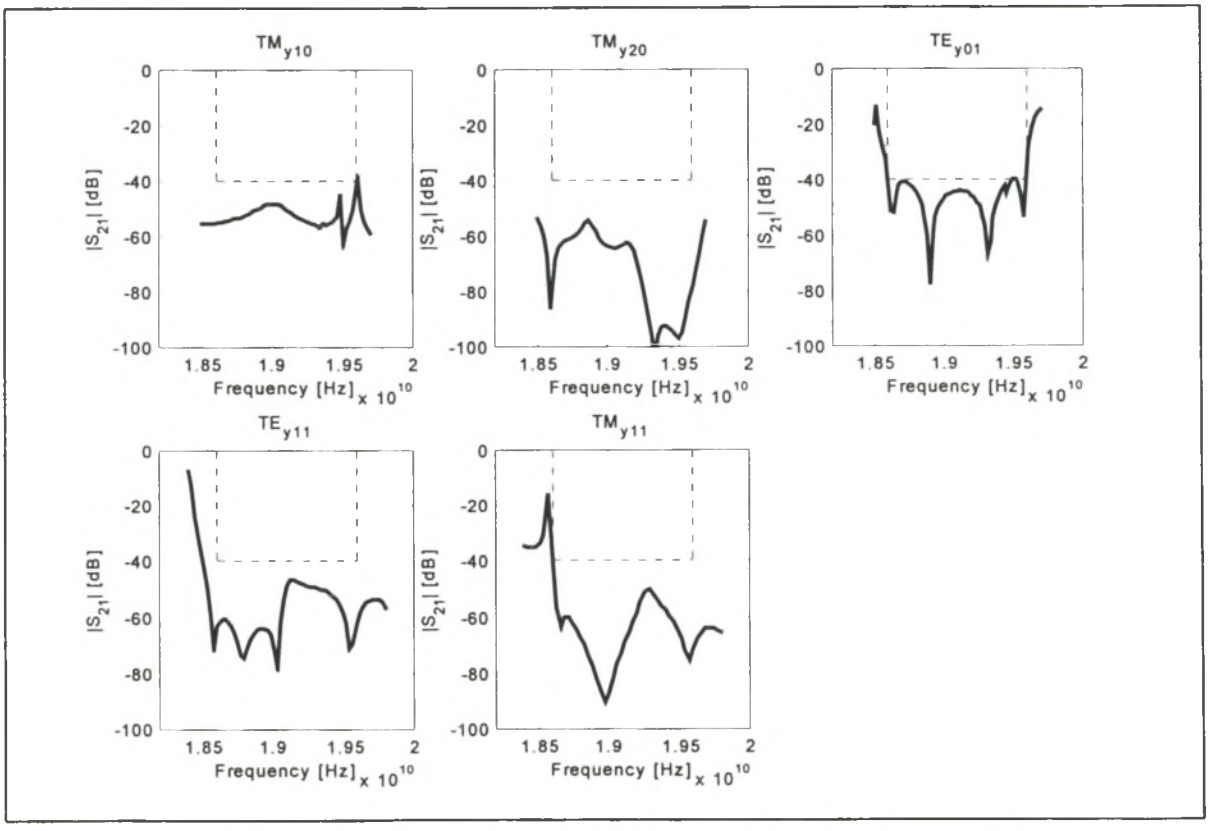


Figure 5.12 Simulated $S_{21}(m,m)$, for m = 1..5 (propagating modes)

Figure 5.13 is a simulated graph of the performance as measured in terms of the ratio of total output power to total input power for a number of random incident modal excitations, shown in a Monte Carlo plot. This analysis assumes that incident power will be distributed in an unknown fashion between the modes - the worst-case distribution provides an upper boundary of the plot and gives an idea of the maximum power that will be allowed to propagate through the filter under the worst conditions.

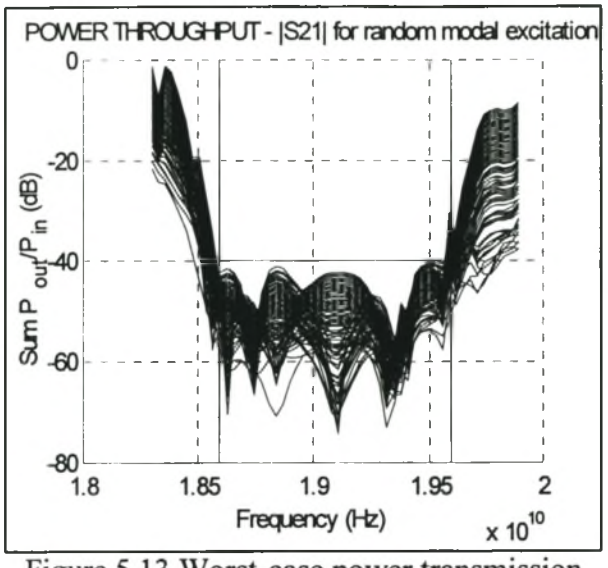


Figure 5.13 Worst-case power transmission

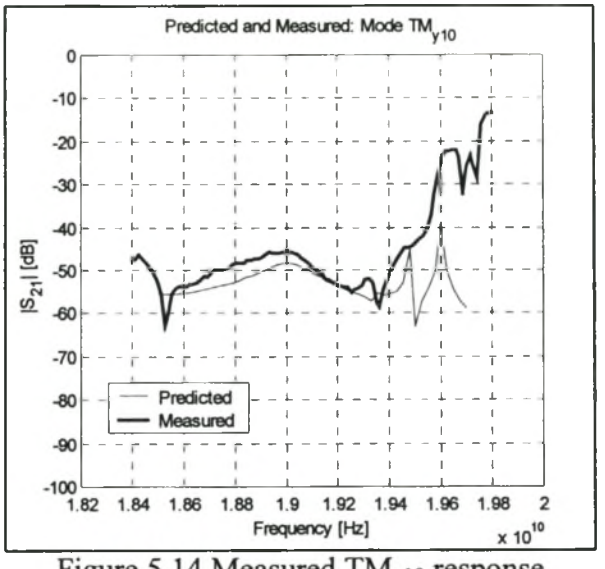


Figure 5.14 Measured TM_{y10} response

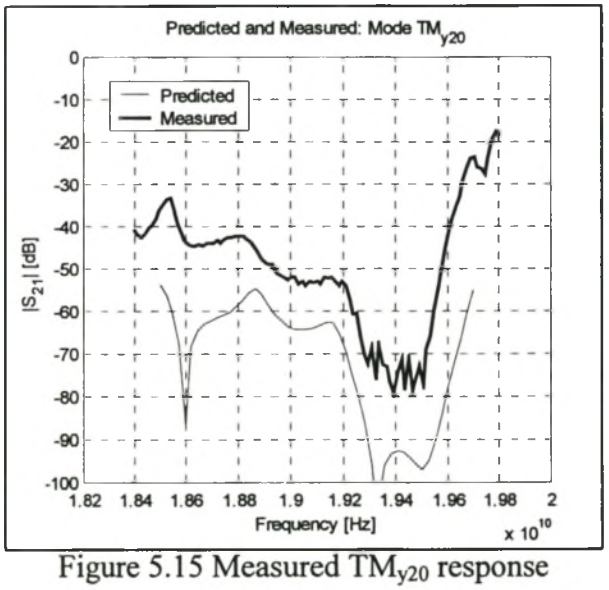


Figure 5.15 Measured TM_{y20} response

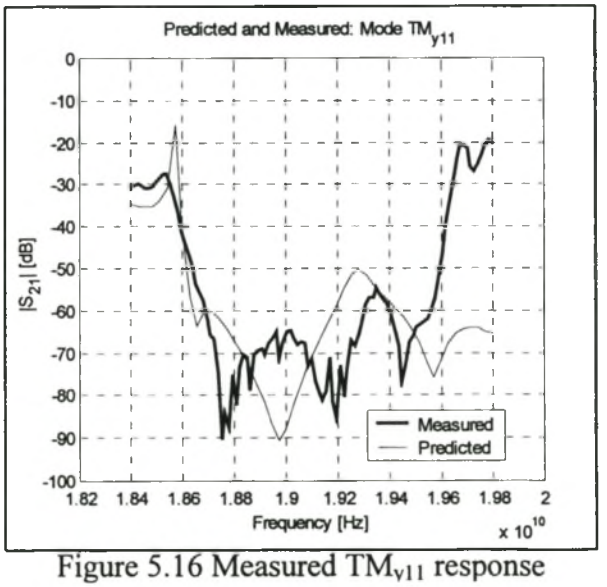


Figure 5.16 Measured TM_{y11} response

The filter was milled in aluminium and measured using the techniques described in chapter 4. Figures 5.14 to 5.16 compare the measured results of the filter to those predicted from the mode matching technique for the TM_{y10}, TM_{y20} and TM_{y11} modes. These three modes are excited by simple probe coupling, and are therefore the easiest to measure. To measure the

other two propagating modes accurately, aperture coupling should be used. It is clear that the transmission parameters are all below 40dB across the band, with good comparison to the predicted results in most cases. The large difference from 19.6GHz upwards is due to a 6^{th} mode starting to propagate. Other deviations from predicted values result from manufacturing and measurement error. Figure 5.17 is a picture of the constructed filter.

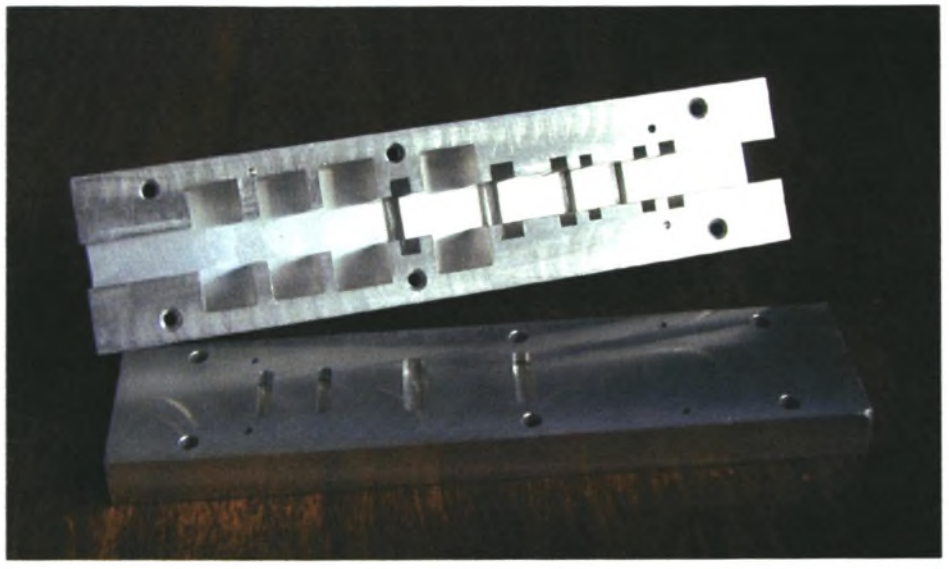


Figure 5.17 The Manufactured 19.1 GHz filter

This filter was originally designed by hand. Subsequent re-designs with the growth-based strategies achieved considerably improved designs over this prototype. Figure 5.18 is a schematic of one of the new designs resulting form the tree-search method. The filter is considerably shorter, 9cm and makes use of only 7 functional blocks compared to 11.5cm and 10 functional blocks in the original design. Figure 5.19 shows the response of the new design.

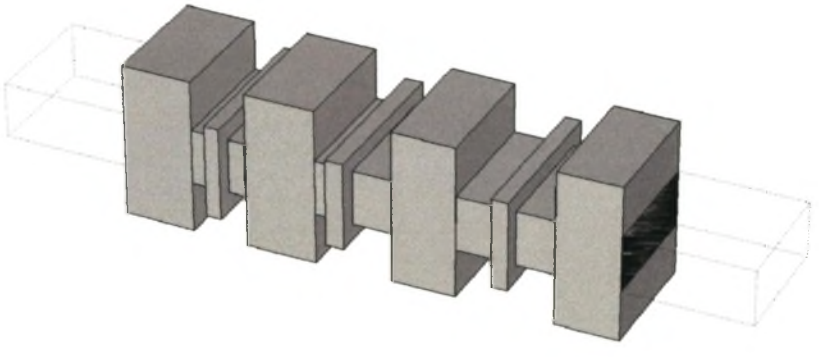


Figure 5.18 Schematic of an improved design of Filter 1
using the Tree-Search Method

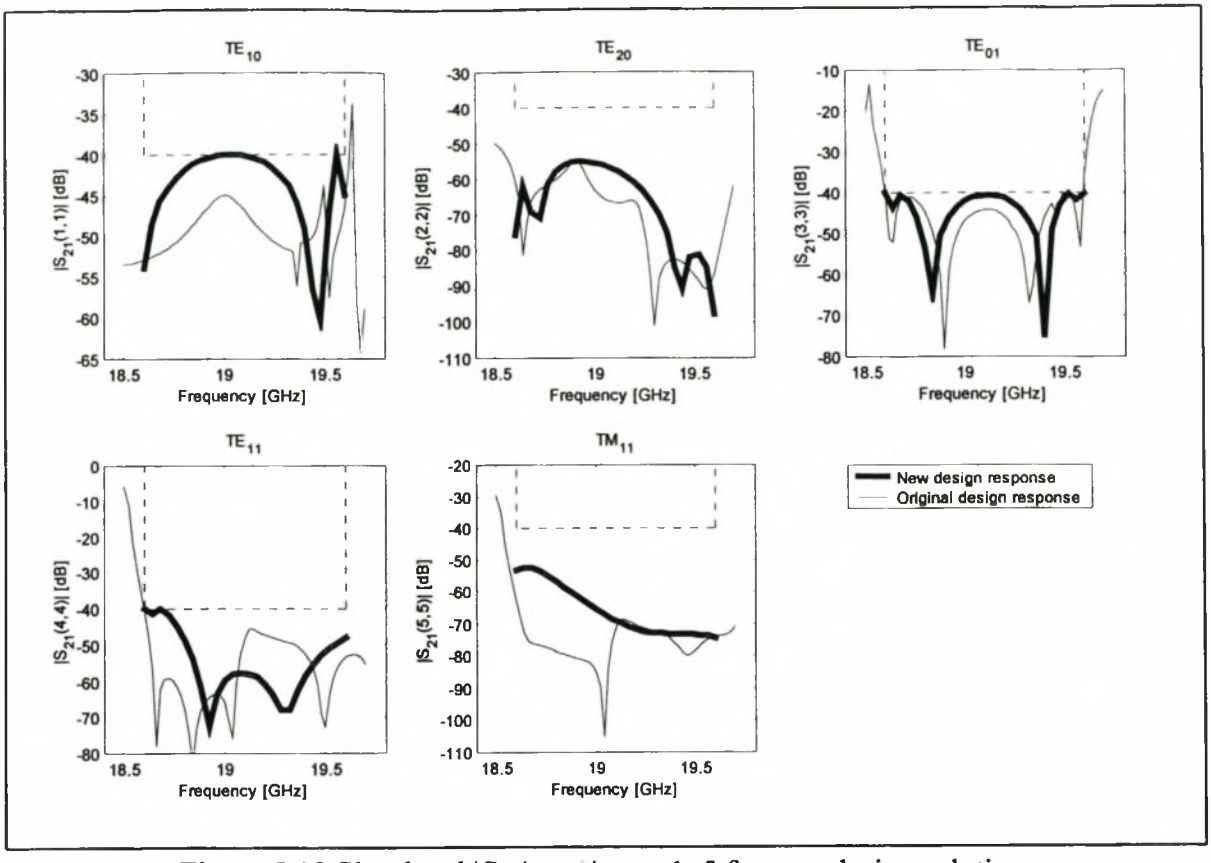


Figure 5.19 Simulated $|S_{21}(m,m)|$, m = 1..5 for new design solution

FILTER 2. (19 GHZ BANDSTOP FILTER – REPEAT FROM CHAPTER 4)

This filter has already been discussed in section 4.6 and is repeated here for the sake of completeness. The centre frequency is 19Ghz, with a bandwidth of 4%. The filter is to provide at least 30dB attenuation in the presence of five propagating modes. Input and output port dimensions are 22.86mm by 10.16mm as before. A schematic of the filter is shown in figure 5.20, with a photograph of the manufactured prototype in figure 5.21. The design was done using the evolutionary strategy.

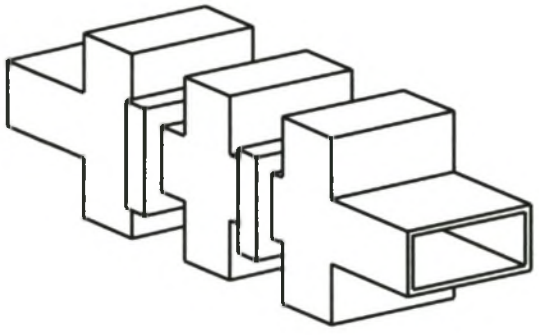


Figure 5.20 Schematic of the 19GHz filter

Figure 5.21 The manufactured filter

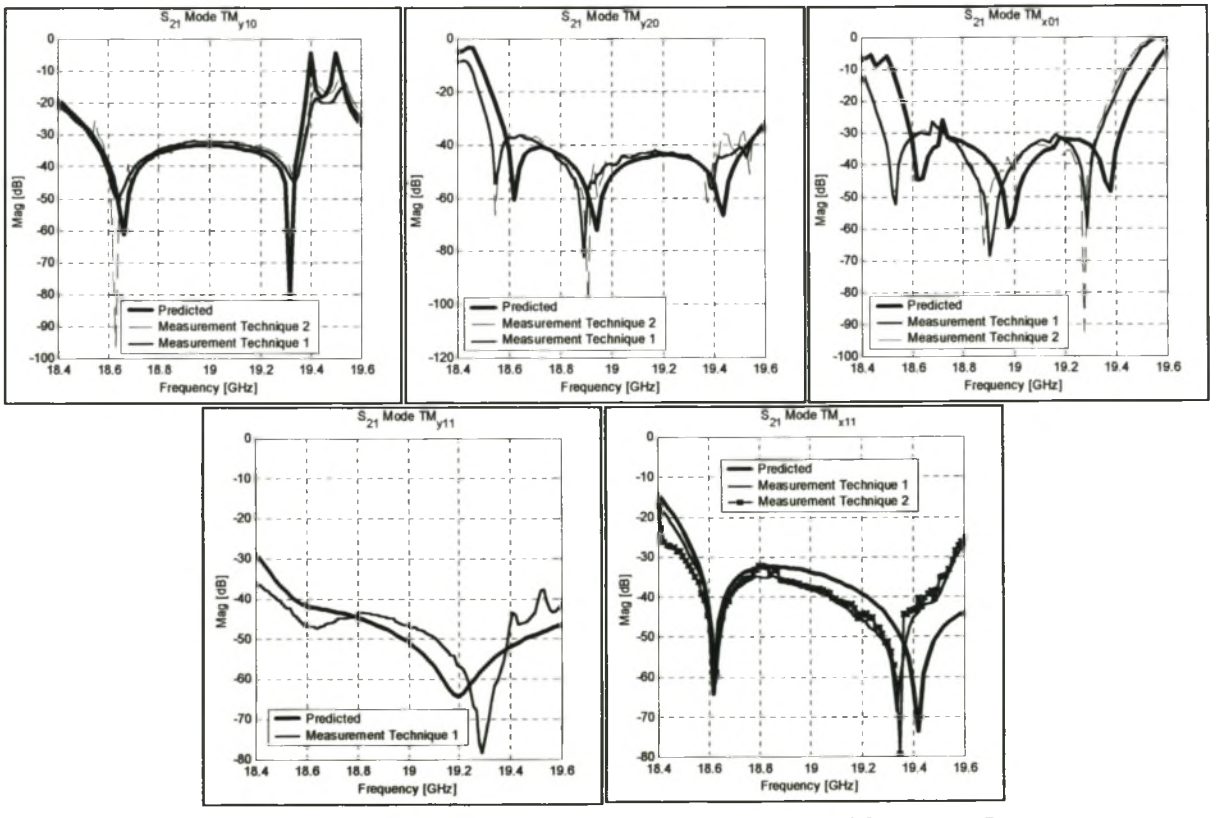


Figure 5.22 Predicted and Measured $|S_{21}(m,m)|$ with $m = 1..5$

The predicted and measured result of the filter are repeated for the sake of completeness in figure 5.22. Slight deviations are due to manufacturing faults of the order of 50 micrometres. The lower dynamic range and improved manufacture of this filter allowed particularly good measurements to be made, verifying the analysis and design principles.

CHOKE 1. (2.45GHZ 30DB MICROWAVE HEATING CHOKE)

This design was required for a practical in-house application. The choke provides 30dB power attenuation to all five propagating modes supported by the input and output port apertures, with dimensions, 160mm by 90mm. The centre frequency is 2.45GHz and the bandwidth is 5%.

A number of design options were generated by the evolutionary strategy. Of these, the shortest, with a length of 0.75m, was selected for manufacture. A schematic and a photograph of the choke are shown in figures 5.26 and 5.27 respectively.

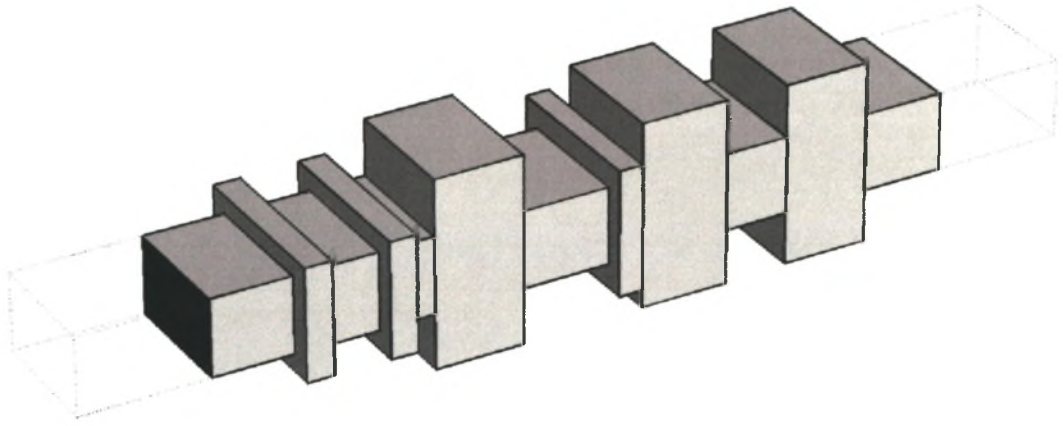


Figure 5.26 Schematic of the Choke

Figure 5.27 Photograph of the choke and its creators

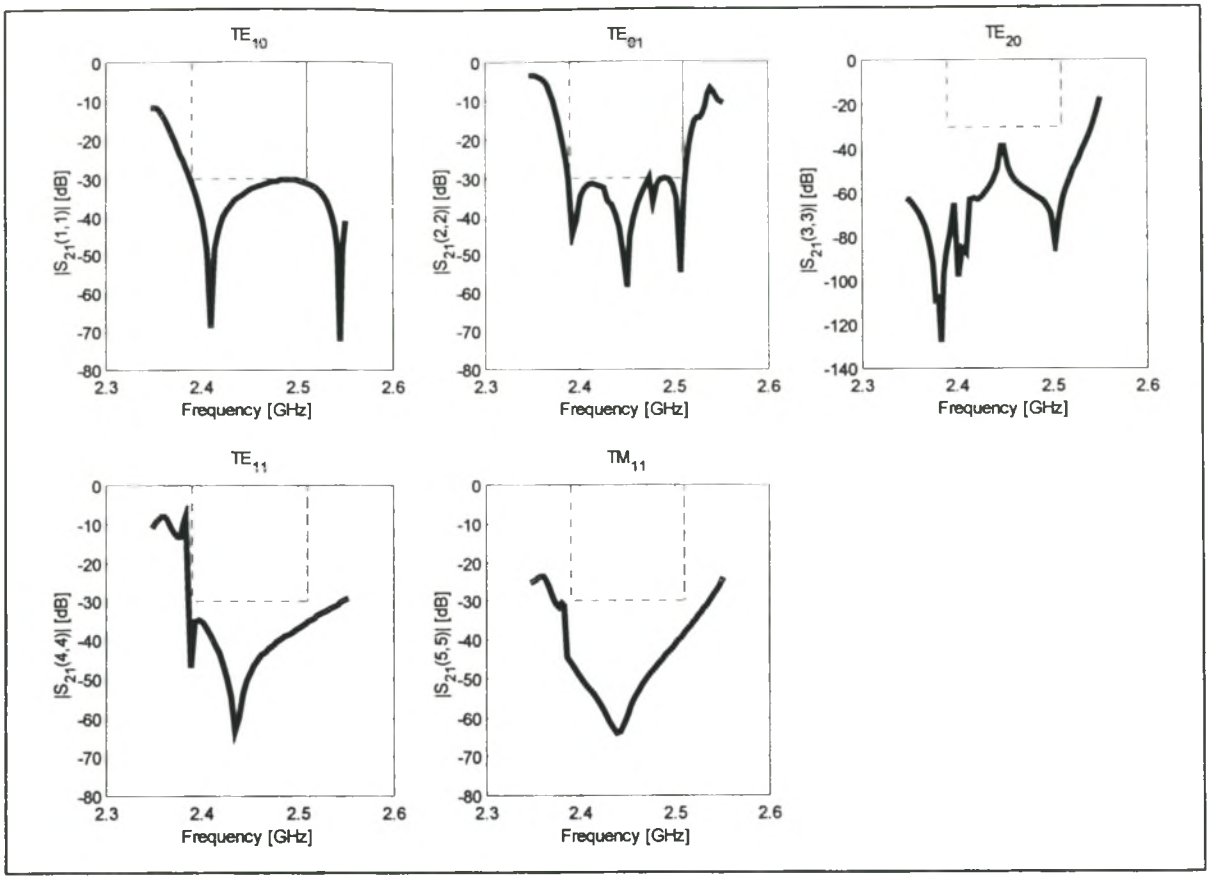


Figure 5.28 Simulated $|S_{21}(m,m)|$, m = 1..5 for 2.45GHz 30dB Choke

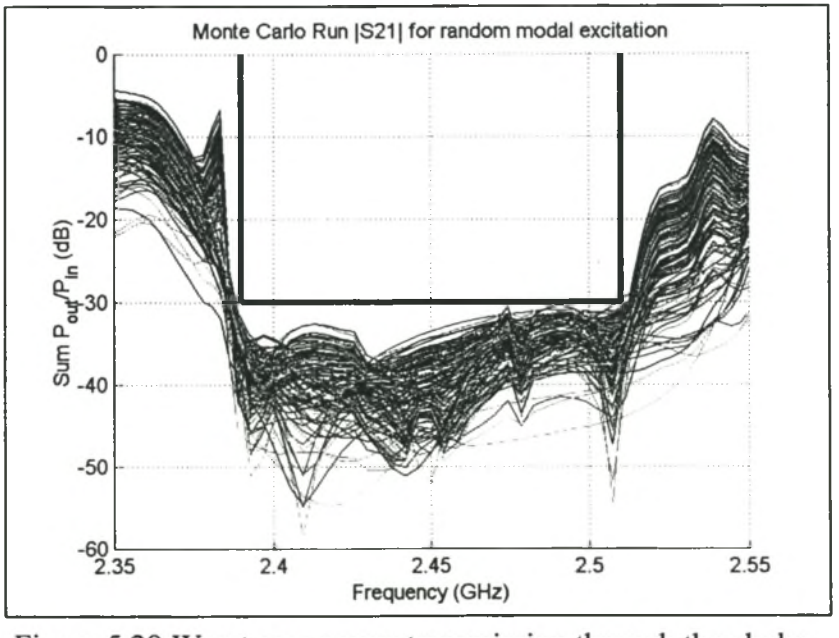


Figure 5.29 Worst case power transmission through the choke

The size and mounting arrangement of the choke made it difficult to measure accurately. Probe-based measurements using only 'through' calibration indicated attenuation from inside the heating cavity to the output port of the choke of at about 40dB. As predicted by various

simulations, this increase over the design value is expected due to the step interface between the heating cavity and the choke aperture (see figure 5.1). The choice of attenuation and bandwidth ensured that power density outside the choke would remain below acceptable safety standards. Further tests with the choke semi-loaded with damp wood did not result in unsafe power levels outside the equipment.

5.7 Conclusion

This chapter has introduced a new method of designing overmoded bandstop filters in oversize waveguide. The design uses multiple resonant functional blocks to create propagation nulls for one or more modes at a particular frequency in the band and makes use of growth-based CAD strategies to cascade them to build up the entire desired stop band. Two design strategies have been detailed, and a number of examples of designed and measured filter have been given.

The filters are particularly suited to applications where the interior of the waveguide cannot be obstructed by, for example, pins or septa, such as would be required of a choke for microwave dielectric heating applications. Furthermore, the modular structure of the filter makes it very easy to tune and resistant to localized disturbances of the internal fields, such as might typically occur in a choke application.

Chapter 6 – Design Solutions: Multimode Waveguide Horns

6.1 Introduction

Single mode horn antennas, which radiate only the dominant mode, were first developed in the late 1930s [Barrow 1939]. Examples are the well known pyramidal and conical horns. The simplicity of design, broad band and ease of manufacture of these antennas led to extensive use in a variety of applications, such as radar, communication and metrology. However, as demands on such systems increase, the fundamental limitations of pure mode horns, that of having very little control of the radiation pattern, becomes problematic. There are two applications where control of the radiation pattern is important. The first is in tracking antennas and the second in shaped pattern antennas.

Certain types of tracking radar systems require the design of a 'monopulse antenna' which can distinguish the direction of incident EM waves. In the early 1960s this was typically accomplished by an array of pure mode horns feeding a reflector antenna. This array is termed a 'monopulse feed'. In a landmark paper by Hannan in 1961, the optimum configuration of such an array was investigated [Hannan 1961a]. He proposed that a near optimum configuration might be attained by using multimode horns, which could selectively radiate either TE_{10} or TE_{20} modes [Hannan 1961b]. He does, however, recognize that the practical aspects of designing such a horn might be difficult: "The generation of the proper waveguide modes is certainly a problem of design".

In shaped pattern antenna applications, horns, used either directly as antennas or as feed elements for reflector antennas, often require specific radiation characteristics. These characteristics might be directed toward achieving low sidelobe levels, low cross polarization and/or high aperture efficiency. Since aperture fields dictate the radiation pattern, it has become common practice to make use of multiple modes at the aperture to shape the aperture field in such a way as to radiate a desired field pattern [Olver 1994].

In addition to these traditional applications of multimode horns, a recent development in power combining, termed 'Spatial Power Combining', has called for the development of horn antennas not intended to radiate in space, but rather to distribute energy evenly incident on a grid structure of active devices. An excellent summary of the current state of the art in spatial power combining, due to Harvey et al., indicates that spatial power combining uses free space to couple energy excited by an array of active devices, as shown in figure 6.1 [Harvey 2000].

This strategy provides inherent advantages over conventional transmission line coupled power combiners, which suffer from increasing losses as the number of coupled elements increases, resulting in an upper limit in the total number of possible coupled devices. Harvey et al. discuss various methods of delivery of energy to and from the array and compare them in terms of performance. Bandwidths ranging from 3 to 20% are typically attained, depending on the choice of array topology and method of energy delivery. For effective use of the array, it is essential for the energy to be delivered and transmitted evenly over the surface of the array. A method of addressing this problem is indicated in figure 6.1b, where dielectric loading on the waveguide wall is intended to provide a more even field distribution than would otherwise be attained by the dominant TE_{10} mode alone.

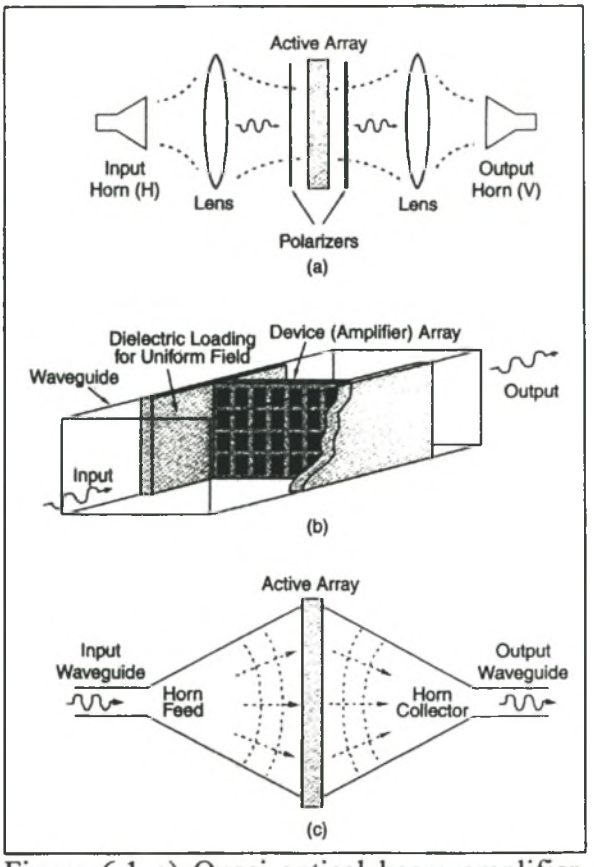


Figure 6.1 a) Quasi-optical beam amplifier concept using a lens-focused arrangement. b) Active array of amplifier elements inside of a waveguide. c) Active array in oversized waveguide with tapered horn feed to accommodate more amplifiers. [Harvey 2000]

The ability to shape the field pattern in the horns of figure 6.1a and 6.1c and the waveguide cross section of figure 6.1b is an important part of the design of spatial power combiners as it dictates the degree to which workload is spread amongst the devices. The use of dielectric

inserts for this purpose can introduce power limitations and losses into such systems that could be avoided by the use of a shaped pattern multimode antenna design.

This chapter proposes a new method of generating multiple mode, shaped aperture fields. The design solutions presented here are applicable to the more traditional shaped pattern and tracking antenna designs, but are more directly aimed at the design of suitable delivery systems for spatial power combiners. The design techniques rely on the CAD principle of structural growth with functional blocks detailed in chapter 2, and represent an original contribution to the field of EM field pattern shaping in multimode apertures.

6.2 A Review of Multimode Horn Designs

The use of multiple modes to shape the field patterns of radiating structures occurred as early as 1962, with the diagonal horn antenna [Love 1962]. This used a combination of TE_{10} and TE_{01} modes in a square aperture with the effect of superior sidelobe and beamwidth performance over conventional pyramidal horns. The very well known and widely used Potter Horn, proposed in 1963, accomplished similar gains for conical horns by mixing the higher order circular TM_{11} mode with the dominant TE_{11} mode [Potter 1963]. Potter divided the design into two stages; mode generation, and mode phasing. Mode generation used an axially symmetric circular step to excite the modes to the correct ratio, followed by a mode suppression section that excluded unwanted modes. The design of the step and suppression section was accomplished empirically. The desired modes were brought into phase by selecting a suitable length of conical horn to follow the mode generation section. Finally, the horn was matched with a circular iris at the mouth. Potter's design principle would serve as the basis for a large number of multimode horn designs in both shaped pattern and tracking antenna applications for many years to come.

The advent of the corrugated conical horn in the early 1970's saw a strong growth in the use of circular aperture horns. These horns employed hybrid modes, excited by the periodic corrugations, to produce excellent electrical performance. The cost was increased weight and complex manufacture. In 1981, Clarricoats et al. combined a multimode feed with a corrugated circular horn for a tracking radar application [Clarricoats 1981]. This configuration

provided improvements in aperture efficiency and low cross polarization. Their design technique was essentially empirical in nature, incorporating a parametric analysis which provided some 'rule of thumb' values to aid the design.

Later years also saw improvements on the well known single step transducer used in Potter's design. Dielectric rod, tube and metal ring inserts were used to excite the desired aperture modes in an effort to improve the limited bandwidth of the step transducer and to provide reductions in cross polarization. A good summary of these efforts is contained in [Lier 1986]. This paper indicates that the Potter-style design is limited to a theoretical maximum bandwidth of about 2.5%, and details design formulas for multimode horns using alternative transducers that provide low cross polarization for bandwidths of up to 3.05%. Lier also emphasizes the growing importance of space based communications and its requirement for low sidelobe power and cross polarization.

Two years later, Lee et al. presented the design and analysis of the multimode rectangular feed horn proposed by Hannan [Lee 1988]. As with the Potter horn, their horn consists of three components; a matching section, a multimode step transducer and a horn flare, with the added complication of having to simultaneously provide for both the TE_{10} and TE_{20} mode excitation. Their design approach also mirrored Potter's in that the individual components were designed separately with the aid of empirical techniques and a quantitative analysis of the effect of various physical variables. Their horn provided up to a 10% useful bandwidth with a return loss of the order of -17dB.

In 1989, Wriedt et al. indicated a growing interest in rectangular apertures, which had some advantages over the more widely used circular apertures, made attractive by the Potter and corrugated horns [Wriedt 1989]. In addition to the convenience of the Cartesian coordinate system, rectangular apertures show greater flexibility in generating a wide variety of beam characteristics, an easier fabrication and a simpler integration with other (traditionally rectangular) waveguide components. This paper extends the Potter style design by employing an evolutionary strategy to aid the design of the transducer section at the mouth of the horn. This strategy optimises the values of a fixed number of step heights and separating lengths according to a predefined error function. A rigorous analysis of the entire radiating structure is accomplished by combining mode matching for the horn structure and the method of

moments for the radiating aperture. Their technique is employed in the design of an enhanced rectangular aperture horn with improved pattern symmetry. The design employs a double step transducer section with a standard flare and accomplishes a bandwidth of about 3.63% (where improved pattern symmetry is observed). The return loss of the device is, unfortunately, not provided. A second design improves the cross polarization characteristic of the first design, which was –22dB, by optimising the transducer and flare to provide for minimum cross-polarization. This does, however, result in lower pattern symmetry.

The application of horns to spatial power combining is detailed by Ivanov et al. in 1996. They make use of a standard hard horn lined with a dielectric slab and incorporating conducting strips to improve the uniformity of the aperture field. The device achieves approximately a 3% bandwidth at 10 GHz with a typical return loss of about 15dB [Ivanov 1996]. Later, Ali et al. detailed the analysis and measurement of hard horn feeds incorporating dielectric slab-linings and using a lens to correct for phase variation over the aperture [Ali 1998]. They achieve a ±1 dB variation of aperture field over more than 98% of the aperture at the design frequency of 31GHz. Most recently, Ortiz et al. presented a design for a 25W and 50W Ka-band quasi-optical amplifier using hard horn feeds [Ortiz 2000]. As before, this design makes of dielectric slabs lining the horn walls in the E-plane to accomplish a uniform aperture field distribution. The horns achieve a -3dB back-to-back return loss with approximately a 3% - 10dB insertion loss. The design made use of a dielectric lens to compensate for aperture phase variation.

One of the most recent contributions to the design of multimode horns is due to Granet et al. [Granet 2000]. This details the design of a new compact multimode corrugated horn with low sidelobes, specifically for satellite applications. The new multihybrid mode horn uses the Potter design philosophy, including a multihybrid mode generator section and a flared section that brings the multiple hybrid modes into phase. The hybrid modes are supported by corrugations along the entire length of the horn. The profile of the flare section is described by a sine function whose parameters are determined empirically. The mode generator section consists of a stepped section of corrugated horn, resembling the Potter configuration. Of particular interest is the importance of compactness highlighted by the authors. Their horn achieves a bandwidth of about 2.5% with an overall length of $5.6\lambda_0$, which is considered compact compared to typical designs of length $30\lambda_0$. The aperture diameter is $4.6\lambda_0$.

It is generally accepted that the drawbacks of multimode horns are limited bandwidth and complicated design due to a number of viable geometries. There are relatively few general designs, and typically an optimal design must be found for each new set of constraints [Olver 1994]. In spatial power combining applications, extensive use is made of dielectric slab linings and lenses to accomplish uniform field distributions. This practice carries with it inherent bandwidth and power level limitations.

Analysis of horns is typically accomplished with the mode matching technique, and if the aperture effects are taken into account, this is usually combined with the method of moments. As yet, general design of such devices follows the principle proposed by Potter, with a matching section, mode transducer and flare section.

6.3 Modal Content for Field Shaping

This section deals with obtaining an appropriate specification of the horn's network parameters, depending on whether it is to be used in an antenna (or reflector antenna feed) application or in a spatial power combiner. The effect of the horn termination will first be discussed. It will be shown that in both applications, the characteristics of primary importance can be approximated as the input side return loss and the field excited at the horn output *under perfectly matched conditions*. Two simple methods of obtaining the desired network parameters given a specified field pattern at the output port of the perfectly matched horn will then be detailed. Finally, an appropriate formulation for the error function of the multimode horn will be given.

6.3.1 Multimode horn terminations

Designs for antenna and feed applications typically specify a desired far field pattern. The far field pattern is predicted from the horn's aperture fields with the use of the Fourier Transform method [Collin 1985]. Various methods of accounting for radiation effects at the aperture, shown in figure 6.2, are discussed in the literature. In many cases it is assumed that for large apertures, of the order of $3\lambda_0$ or larger, there will be negligible reflection and cross coupling at

the aperture [Kühn 1984], implying that the aperture fields can be considered equal to the incident fields excited by the horn. This approximation is widely supported in the literature [Lee 1988, Lier 1986, Clarricoats 1981]. Other techniques, such as the geometric theory of diffraction [Granet 2000], and the Kirchoff-Huygen Model [Clarricoats 1981] have also been used to account for aperture effects. Some authors pursue a rigorous analysis of the effect of radiation into the half-space with the aid of the moment method applied to integral equation analysis [Bunger 1999] and orthogonal rooftop basis functions [Wriedt 1989].

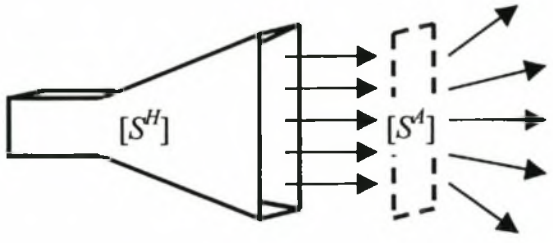


Figure 6.2 The horn and aperture in an antenna application

In spatial power combining systems, where the horn's output port is directly coupled to the active array, there is no radiation in the half space, and consequently, no aperture effects. Disruptions of field at the horn output port may, however, occur due to a mismatch at the active array. This is, unfortunately, dependent on the characteristics of the specific array and cannot be accounted for in this dissertation.

For the purposes of this investigation, a perfect match will therefore be assumed between the output port of the horn and the terminating radiating aperture (for antenna applications) or grid of active devices (for spatial power combining applications). If so desired, a more rigorous approach could incorporate accurate analysis of aperture effects or models of the active array without limiting the effectiveness of the design procedures presented here.

6.3.2 Determining modal content for a specified field pattern

Under perfectly matched conditions, the EM fields excited at the output port of the horn can be approximated as in eqns. (6.1), where $[S^H]$ is the GSM of the horn with M modes at the output port, and $\mathbf{E}^t_{1..M}$ and $\mathbf{H}^t_{1..M}$ are the normalized tangential modal fields at the output port as indicated in eqns. (3.5) and (3.6).

$$\begin{aligned}\mathbf{E}^t_{\text{Tot}}(x,y) &\approx \left[\mathbf{E}^t_1(x,y) \quad \mathbf{E}^t_2(x,y) \quad \ldots \quad \mathbf{E}^t_M(x,y)\right]\cdot \bar{S}^H_{BA}\cdot a_A \\ \mathbf{H}^t_{\text{Tot}}(x,y) &\approx \left[\mathbf{H}^t_1(x,y) \quad \mathbf{H}^t_2(x,y) \quad \ldots \quad \mathbf{H}^t_M(x,y)\right]\cdot \bar{S}^H_{BA}\cdot a_A\end{aligned} \tag{6.1}$$

Note that this is an approximation as the number of modes considered at the output port have been truncated to a finite value. Typically, M would be the number of propagating modes at the output port and eqns. (6.1) would be consistent only if either a short length of uniform line were incorporated at the output port to sufficiently attenuate any non-propagating modes, or if there was negligible excitation of non-propagating modes near the output port. Evaluation of the required horn transmission parameters, $\bar{S}^H_{BA}$, can be accomplished in two ways.

a) Use of modal orthogonality

This method is appropriate if either the output port **E**- or **H**-field are exactly specified. It uses orthogonality between the waveguide modes to determine the required relationship between the modes at the output port in terms of phase and magnitude. Assuming equality between the left and right hand sides of eqns. (6.1), we have eqns. (6.2) for an **E**-field specification and similarly, eqn. (6.3) for an **H**-field specification. $\tilde{e}_m$ and $\tilde{h}_m$ are the modified modal field patterns for mode, m, as defined in eqns. (3.5) and (3.6) and Z_m and Y_m are the impedance and admittance of mode m.

$$\begin{aligned}&\iint_{S_{WG}}\left[\left(\mathbf{E}^t_{\text{Tot}}(x,y)\times \tilde{h}_m(x,y)\right)\cdot \hat{i}_z\right]\cdot dA = \sqrt{Z_m}\cdot S^H_{BA}(m)\cdot a_A \\ &\therefore a_A\cdot S^H_{BA}(m) = \varsigma^H_{BA}(m) = \frac{\iint_{S_{WG}}\left[\left(\mathbf{E}^t_{\text{Tot}}(x,y)\times \tilde{h}_m(x,y)\right)\cdot \hat{i}_z\right]\cdot dA}{\sqrt{Z_m}}\end{aligned} \tag{6.2}$$

$$a_A\cdot S^H_{BA}(m) = \varsigma^H_{BA}(m) = \frac{\iint_{S_{WG}}\left[\left(\tilde{e}_m(x,y)\times \mathbf{H}^t_{\text{Tot}}(x,y)\right)\cdot \hat{i}_z\right]\cdot dA}{\sqrt{Y_m}} \tag{6.3}$$

Note that the value of excitation, a_A, is arbitrary, and must be chosen so that $\bar{S}^H_{BA}$ is physically realizable. The error function used in the design, however, expresses the transmission parameters in terms of relative magnitude and phase, and uses the values of $\bar{\varsigma}^H_{BA}$.

b) Use of optimisation

Optimisation allows a more flexible approach to determining the values of $\bar{\varsigma}_{BA}^{H}$ to be used in the design. It also allows simultaneous expression of desired **E**- and **H**- field patterns and obtains the best compromise in transmission parameters to achieve these patterns. In this approach the magnitudes and phases of the entries of $\bar{\varsigma}_{BA}^{H}$ are selected, with the aid of an iterative least squares solver, to minimize the difference between the specified field pattern(s) and those obtained by eqn. (6.1). Note that this approach can also be used to solve for the appropriate modal weights to satisfy a specified *complex power* distribution by forming the cross product of the two quantities in eqn. (6.1).

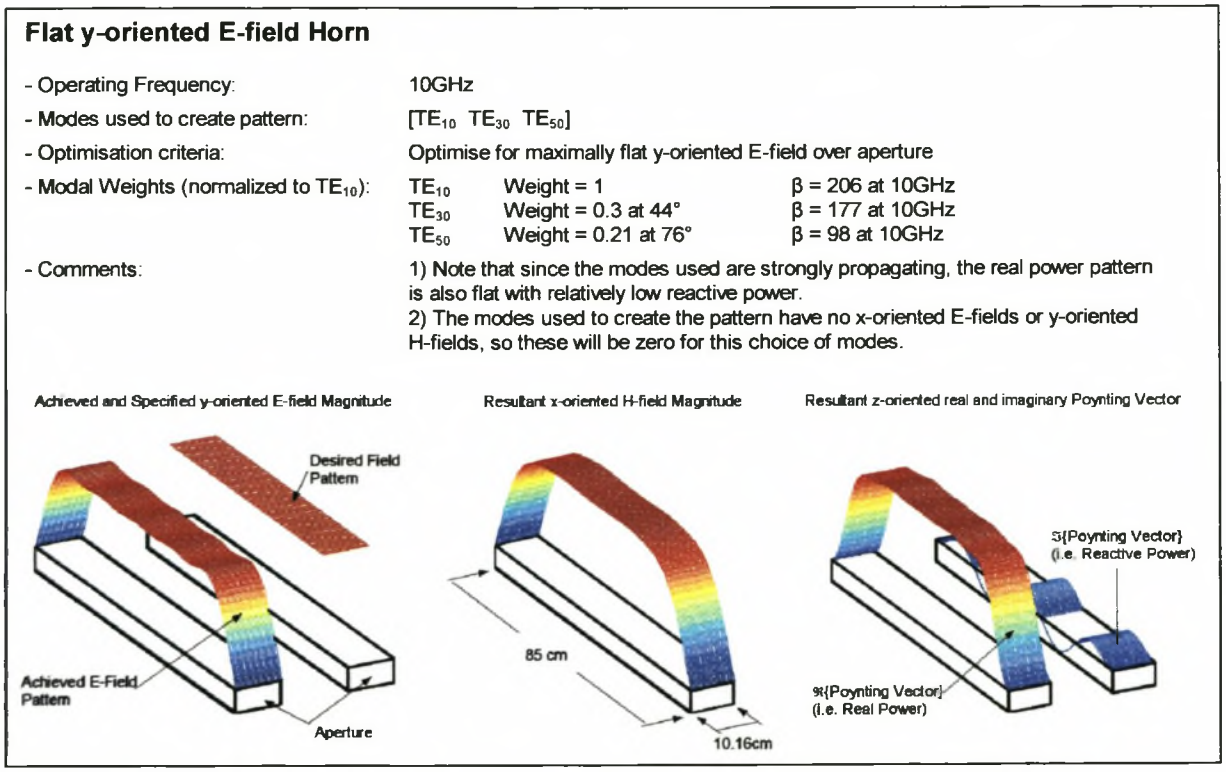


Figure 6.3 Example of field pattern shaping in an H-plane horn application

Figure 6.3 shows sample field patterns synthesized with this approach and the corresponding modal weights, $\bar{\varsigma}_{BA}^{H}$, for a horn application. The modes used to build up the patterns were selected from those modes that were propagating in the waveguide at the design frequency. The use of strongly propagating modes in the example above ensures that the **H**-field and the real power distribution are both flat. Weakly propagating modes tend to have wave impedances further from the intrinsic impedance, and consequently have different relative

quantities of **E** and **H** fields. If a horn were to be designed for an optimally flat **E**-field pattern using such modes, the **H**-field pattern, and consequently, the power distribution could not be equally flat. A general rule is that designs that wish to show both flat **E** and **H** field patterns must use modes with similar wave impedances. Figure 6.4 shows the result of incorporation of more TE_{m0} modes to further flatten the **E**-field pattern. Because some of the new modes are below cut-off there is now a non-uniform real power distribution at the aperture, but a much flatter **E**-field pattern.

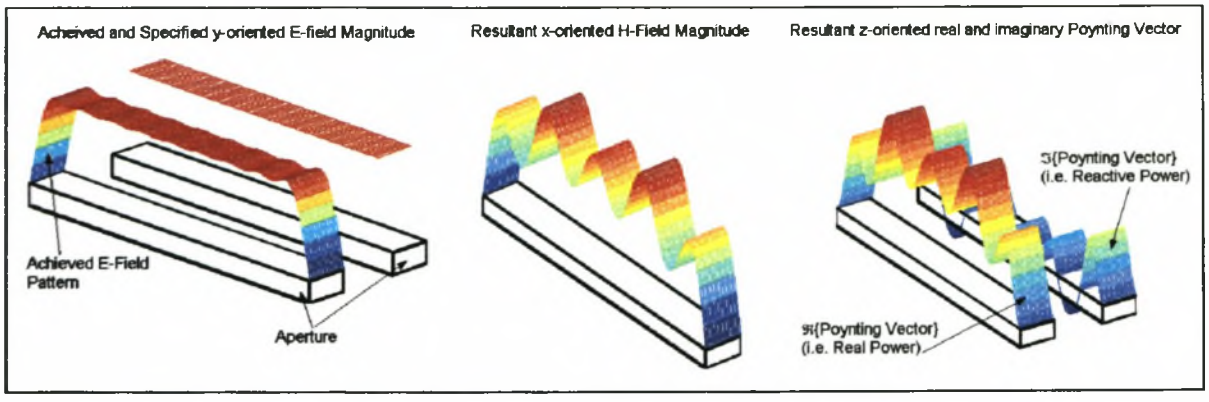


Figure 6.4 Incorporation of TE_{70} and TE_{90} modes for increased **E**-field uniformity

While it is possible to make use of non-propagating modes to form a field pattern, care must be taken to ensure that if the output port length is specified (i.e. as a uniform length of waveguide before the reference plane), that the modes are not excessively attenuated by this line, resulting in a near impossible problem for the design strategy, which must ensure that appropriate relative quantities of these modes reach the reference plane. For these reasons, it is preferred to make use of propagating modes in the examples detailed in this chapter.

When the modal weights are obtained in the form of $\bar{\varsigma}_{BA}^{H}$, the error function of the desired horn can be specified in terms of these relative weights.

6.3.3 Multimode horn error functions

In some applications it may be necessary to specify exact magnitudes, phases or both or certain GSM entries. Error functions suitable for this kind of problem can generally be specified according to the guidelines in chapter 2. In the special case of multimode horns

designed to realize a specified field pattern, however, the error set specified below in eqns. (6.4) to (6.7), with the weight ratios, ρ_{kQ}^{X}, given in eqns. (6.8), is especially appropriate. The error set consists of three parts. First is the return loss over frequency, shown in eqn. (6.4), which is simply to be minimized. Second is the difference in relative magnitude of the modes excited at the output port over frequency. The last part is the difference in relative phase of the modes to be excited at the output port. The different error types are combined in the final error set in eqn. (6.7) with weighting coefficients, *w*, which can be set according to the relative importance of the different error types. In this investigation, the weights were chosen as in eqns. (6.9).

Reflection Component:

$$\bar{e}^{Refl} = \left[\left\{ S_{AA}^{H}(f_1) - \varsigma_{AA}^{H}(f_1) \right\} \quad \left\{ S_{AA}^{H}(f_2) - \varsigma_{AA}^{H}(f_2) \right\} \quad \cdots \quad \left\{ S_{AA}^{H}(f_N) - \varsigma_{AA}^{H}(f_N) \right\} \right]$$

with (6.4)

$$\varsigma_{AA}^{H}(f_n) = 0 \quad \forall \quad n = 1..N$$

Magnitude Component:

$$\bar{e}^{Mag} = \left[\bar{e}_1^{Mag} \quad \bar{e}_2^{Mag} \quad \cdots \quad \bar{e}_k^{Mag} \quad \cdots \quad \bar{e}_M^{Mag} \right] \qquad \forall \quad k = 1..M, \; k \neq Q$$

with (6.5)

$$\bar{e}_k^{Mag} = \left[\left\{ \left| \left| \rho_{kQ}^{H}(f_1) \right| - \left| \rho_{kQ}^{D}(f_1) \right| \right| \right\} \quad \left\{ \left| \left| \rho_{kQ}^{H}(f_2) \right| - \left| \rho_{kQ}^{D}(f_2) \right| \right| \right\} \quad \cdots \quad \left\{ \left| \left| \rho_{kQ}^{H}(f_N) \right| - \left| \rho_{kQ}^{D}(f_N) \right| \right| \right\} \right]$$

Angle Component:

$$\bar{e}^{Ang} = \left[\bar{e}_1^{Ang} \quad \bar{e}_2^{Ang} \quad \cdots \quad \bar{e}_k^{Ang} \quad \cdots \quad \bar{e}_M^{Ang} \right] \qquad \forall \quad k = 1..M, \; k \neq Q$$

with (6.6)

$$\bar{e}_k^{Ang} = \left[\left\{ \left| \sin\left(\frac{\angle \rho_{kQ}^{H}(f_1) - \angle \rho_{kQ}^{D}(f_1)}{2} \right) \right| \right\} \quad \left\{ \left| \sin\left(\frac{\angle \rho_{kQ}^{H}(f_2) - \angle \rho_{kQ}^{D}(f_2)}{2} \right) \right| \right\} \quad \cdots \quad \left\{ \left| \sin\left(\frac{\angle \rho_{kQ}^{H}(f_N) - \angle \rho_{kQ}^{D}(f_N)}{2} \right) \right| \right\} \right]$$

Complete Error Set:

$$\bar{e} = \left[w^{Refl} \cdot \bar{e}^{Refl} \quad w^{Mag} \cdot \bar{e}^{Mag} \quad w^{Ang} \cdot \bar{e}^{Ang} \right]^T \tag{6.7}$$

Modal Weight Ratios:

$$\begin{aligned} \rho_{kQ}^{D}(f_i) &= \frac{\varsigma_{BA}^{H}(k, f_i)}{\varsigma_{BA}^{H}(Q, f_i)} \qquad \forall\, k = 1..M,\ k \neq Q \quad \text{(The desired weight ratios)} \\ \rho_{kQ}^{H}(f_i) &= \frac{S_{BA}^{H}(k, f_i)}{S_{BA}^{H}(Q, f_i)} \qquad \forall\, k = 1..M,\ k \neq Q \quad \text{(The actual weight ratios)} \end{aligned} \tag{6.8}$$

$$w^{Refl} = 1 \qquad w^{Mag} = 3 \qquad w^{Ang} = 2 \tag{6.9}$$

If the horn is to perform an additional filtering action, this can be accounted for by specifying a return loss function over frequency in ς_{AA}^{H}.

6.4 Functional Block Selection

As detailed in chapter 2, design is made significantly more efficient by the selection of a limited number appropriate functional blocks to build the structure. In the case of multimode horns, functional block selection is guided by a number of observations from the literature.

- There is typically a demand for low return loss.
- Resonant behaviour (such as in filters) is typically *not* a design objective.
- High useful bandwidth is usually a priority.
- Waveguide steps have been widely used to excite modes in existing horn designs.

Discontinuities such as irises, generally used in filter design, exhibit strong reflections unless employed to introduce resonance, which is not necessarily a desired objective in horn design. Similarly, the resonant structures used in the design of multimode bandstop filters (chapter 5)

intentionally introduce high return losses. It should, however, be noted that these two types of structures could introduce beneficial qualities, such as efficient multimode excitation and selective phase delay if used under the right circumstances (wide, long apertures and shallow, short stubs respectively), and might well be essential for the successful realization of particularly demanding multimode horn designs.

This investigation limits the space of available functional blocks to single- or double-plane waveguide steps in the interests of evaluating the proposed growth-based design strategies. This choice is motivated by the extensive use of these structures in both multimode horn and matching section designs in the literature. In addition, it is proposed that since it is common practice to model the traditionally smoothly tapered flare of the horn as a staircase of waveguide steps [Wriedt 1989], this choice enables the design strategies to recreate the widely-used matching section, mode excitation section, flare section topology should it prove desirable. It is, however, expected that since the design strategies investigated here allow these tasks to be distributed over the entire length of the horn, the emerging finished design might well exhibit a different, potentially more appropriate topology than would otherwise be attained.

Depending on the type of modes to be excited, and the output port geometry, the choice of step type to be used can also be limited. For example, the pattern design in figure 6.3 requires only $TE_{(2n+1)0}$ modes and might suffer from the excitation of other modes. Either prior experience, or use of the equivalent circuit formulation for mode matching (chapter 3), could provide the necessary physical insight to recognize that symmetrical H-plane steps excite only these types of modes, and should therefore be exclusively used for such a pattern design.

6.5 Horn design using growth-based CAD strategy

As opposed to the design problem described in chapter 5, which is primarily oriented toward the arrangement of different functional blocks and well suited to an evolutionary strategy, this problem is solved instead with the scanning strategy. This strategy gradually increases the complexity of the solution by successively obtaining the optimal placement of a single functional block in the structure.

As outlined in chapter 2, an advantage of the scanning strategy is that it is considerably less computationally intensive than the evolutionary strategy. This is critically important in this application, as the functional blocks cannot be solved prior to the design operation as was done with the multimode bandstop filters. Furthermore, as described in chapter 2, the nature of introducing new waveguide steps changes all the physical variables of the system, requiring a full re-analysis of the entire system for each change of a variable. Lastly, since the variation in cross-section between the input and the output waveguides is large with most horn geometries, very large numbers of modes must be used in the mode matching operation to ensure convergence. As a consequence, design of these devices is heavily reliant on the improved analysis techniques detailed in chapter 3. Without the increased analysis speed, the design process could take up to two orders of magnitude longer. Although similar work done by Wriedt et al. [Wriedt 1989] suggests that it would be most certainly possible to use the evolutionary strategies outlined in chapter 2 to accomplish the successful design of these devices, it is estimated that the process would take considerably longer than the scanning strategies to complete a satisfactory design.

The cost of reduced computational effort with the scanning strategy is, however, a reduction in reliability over the evolutionary algorithm. Figure 6.5 shows the yield of the scanning strategy in percent over 60 sample runs of the algorithm. Note that 30% of the designs were unusable, with either unwieldy structures or poor electrical performance. This is in contrast to the very high reliability of the evolutionary strategy as experienced with the designs in chapter 5.

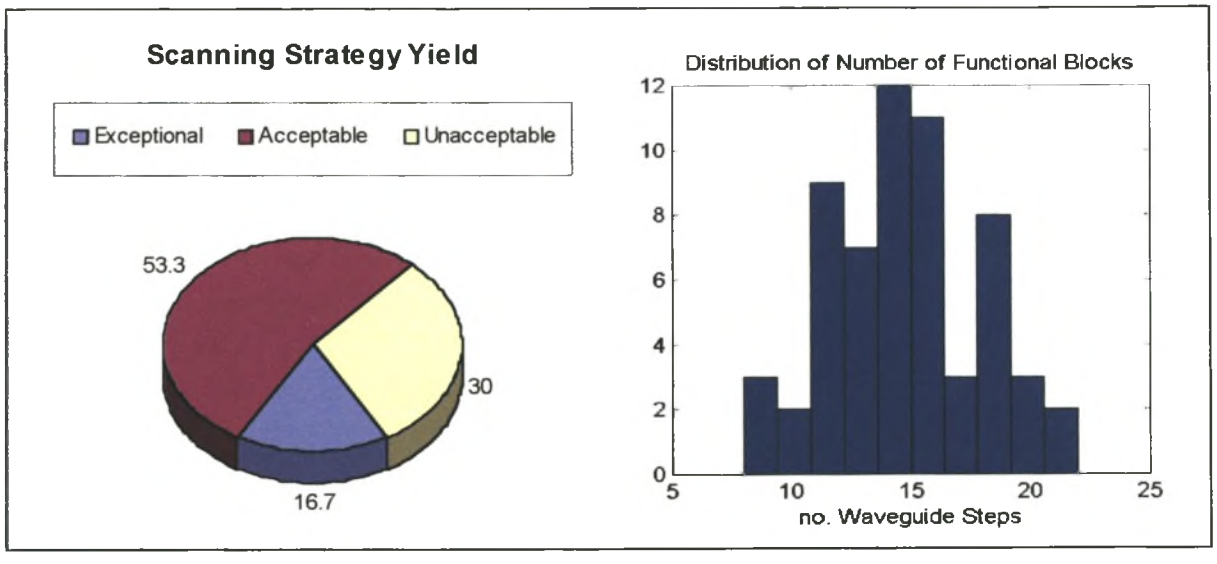


Figure 6.5 Statistical analysis of the behaviour of the Scanning Strategy

The maximum number of functional blocks to include was limited to 22 in this case. The histogram indicates that this value was not generally reached, suggesting that sufficient complexity was attained with as few as 15 blocks. Most of the exceptional designs obtained, used between 13 and 16 blocks. Figure 6.6 shows the development of a sample design of an H-plane horn using the basic Scanning Strategy without checks for redundancy of functional blocks and excessive line lengths. The increase in the structure's complexity along with the gradual decrease in the structure's sum squared error is apparent.

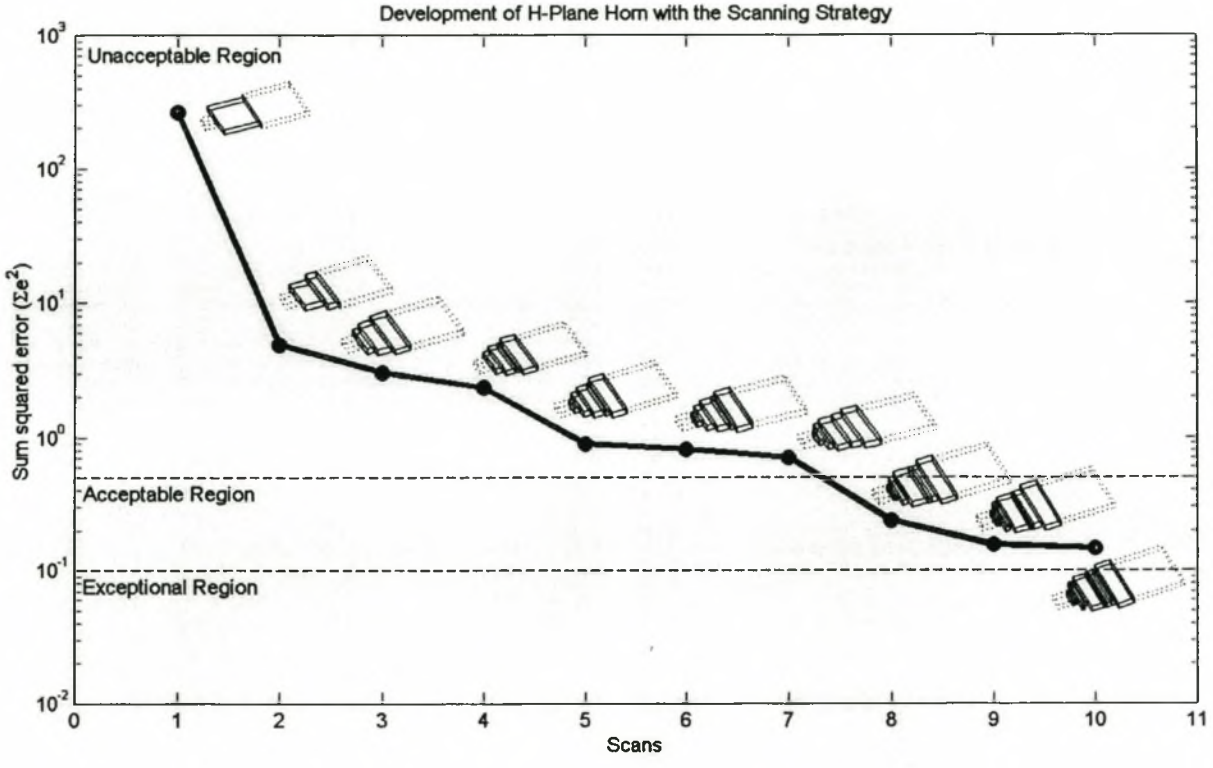


Figure 6.6 The Scanning Strategy in action

In order to verify that the growth-based principle of the scanning strategy provides a meaningful contribution to the design process, a number of attempts were made at direct optimization of a variety of starting horn topologies. As opposed to the growth-based techniques, the number of functional blocks comprising the structure were kept constant at sufficiently high values to ensure a potentially satisfactory design. A large number of varied intelligent first guesses were given to the gradient-based optimization routine and none were observed to converge to satisfactory designs. The best of these designs had sum-squared-error values in excess of 60 times that at the upper limit of acceptability and 6 times that of the very worst solutions ever experienced from the scanning strategy.

6.6 Design Examples

H-PLANE SPATIAL POWER COMBINING FEED (FULL MODE MATCHING)

This horn was designed as a potential feed for an active array with maximally flat E-field and power distribution over the widest possible length of the aperture.

The modal weights acquired to produce this resultant field pattern are as shown in figure 6.3. The desired bandwidth was chosen to be 8%, which lies safely above the 3.5% bandwidth of other existing hard horn designs [Ivanov 1996, Ortiz 2000]. The average design time for this type of problem is about 8 hours on an 866MHz Pentium III. A sketch of the designed horn is shown in figure 6.7.

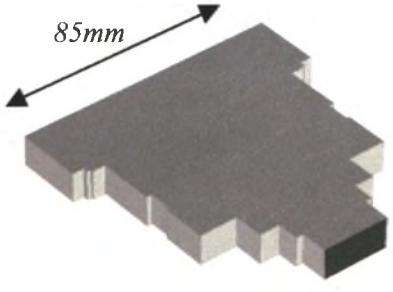


Figure 6.7 Diagram of the H-Plane Feed

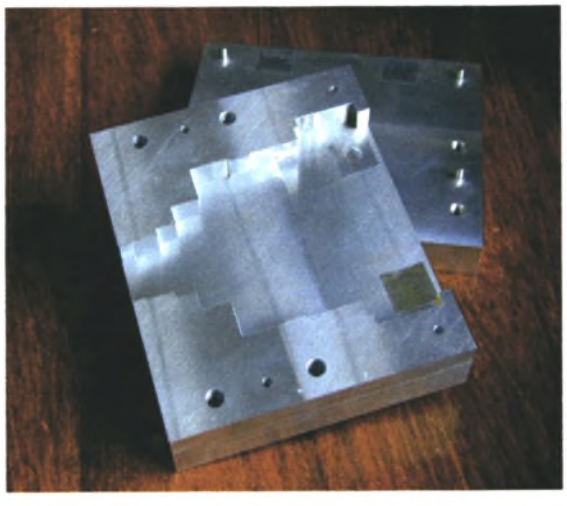


Figure 6.8 The manufactured horn

The horn's predicted insertion loss is shown in figure 6.9. Figure 6.10 shows the optimal modal weight magnitude ratios and phase differences with the corresponding values accomplished by the design procedure. Clearly the design has matched the desired optimal values very closely. Figure 6.11 shows the predicted E-field patterns at the aperture for various frequencies both in, and out of band. Note that the existence of only $TE_{(2n+1)0}$ modes implies that the x-oriented E-fields will be zero and that there will be no variation of y-oriented E-fields in the y-direction, allowing the E-fields to be plotted in the x-dimension alone. A three-dimensional visualization of the fields is, however, shown in figure 6.3.

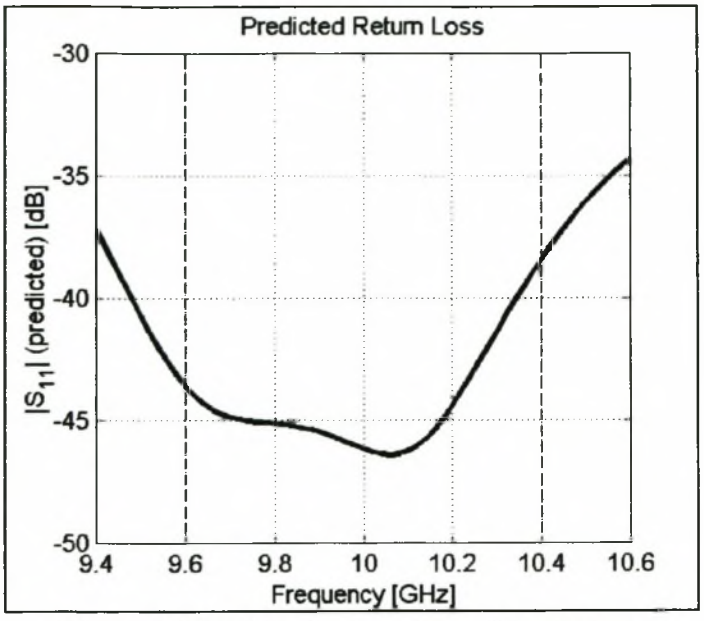


Figure 6.9 Predicted return loss of the design

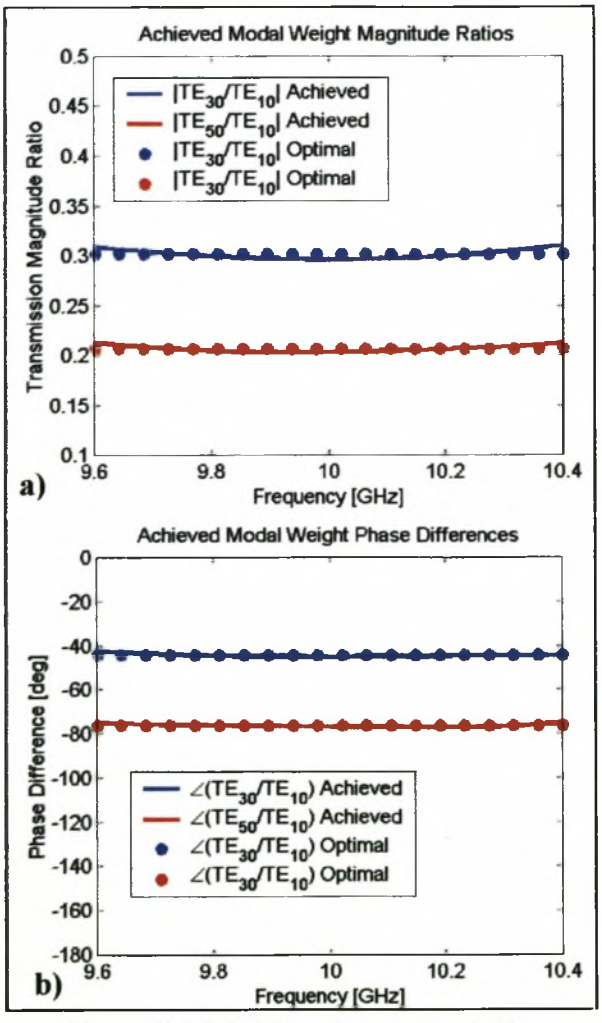


Figure 6.10 Achieved modal weights

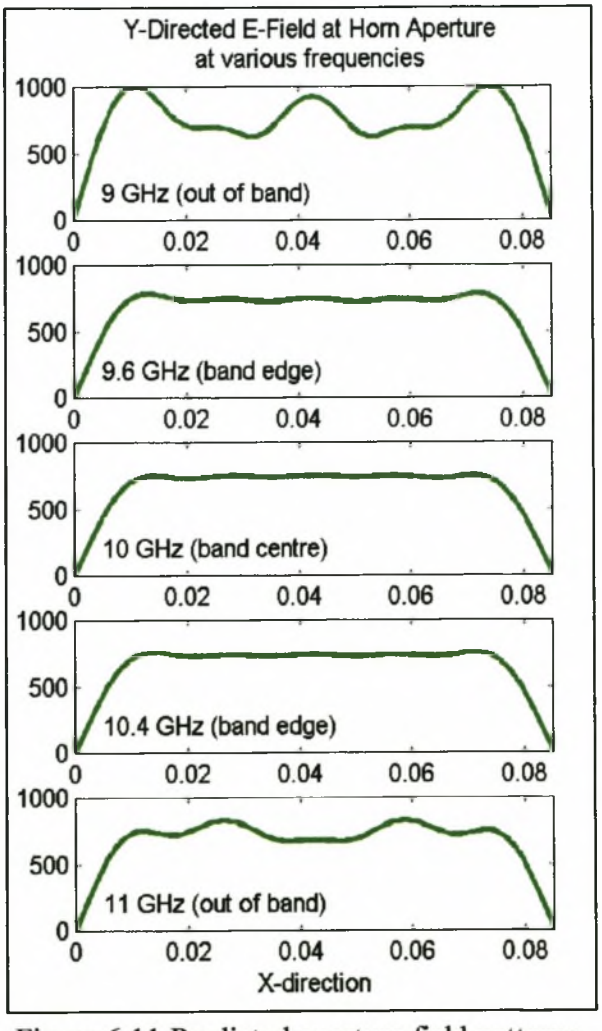


Figure 6.11 Predicted aperture field patterns

The horn was milled from aluminium to an accuracy of about 50μm, and measured with the multimode measurement techniques developed in chapter 4. A photograph of the manufactured device is shown in figure 6.8.

The predicted and measured transmission parameters of the manufactured horn are compared in figures 6.12 and 6.13. The level of noise in the measured results stems from the unavailability of a suitable load for the custom 85x10.16mm waveguide cross-section formed by the aperture of the horn. Simulations of the measurement procedure predicted similar levels of inaccuracy as a result of using a load termination of worse than about -25dB.

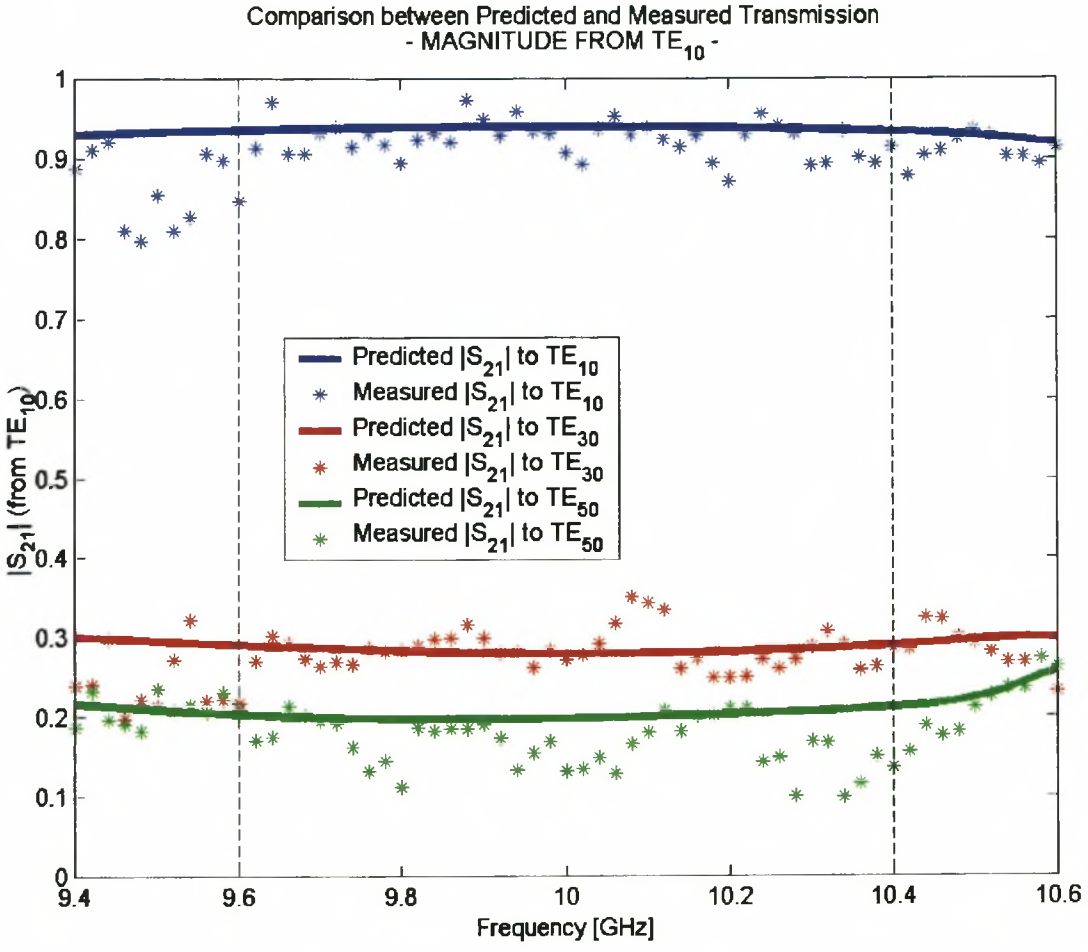


Figure 6.12 Predicted and measured transmission magnitude from TE_{10} mode to the TE_{10}, TE_{30} and TE_{50} modes at the output aperture

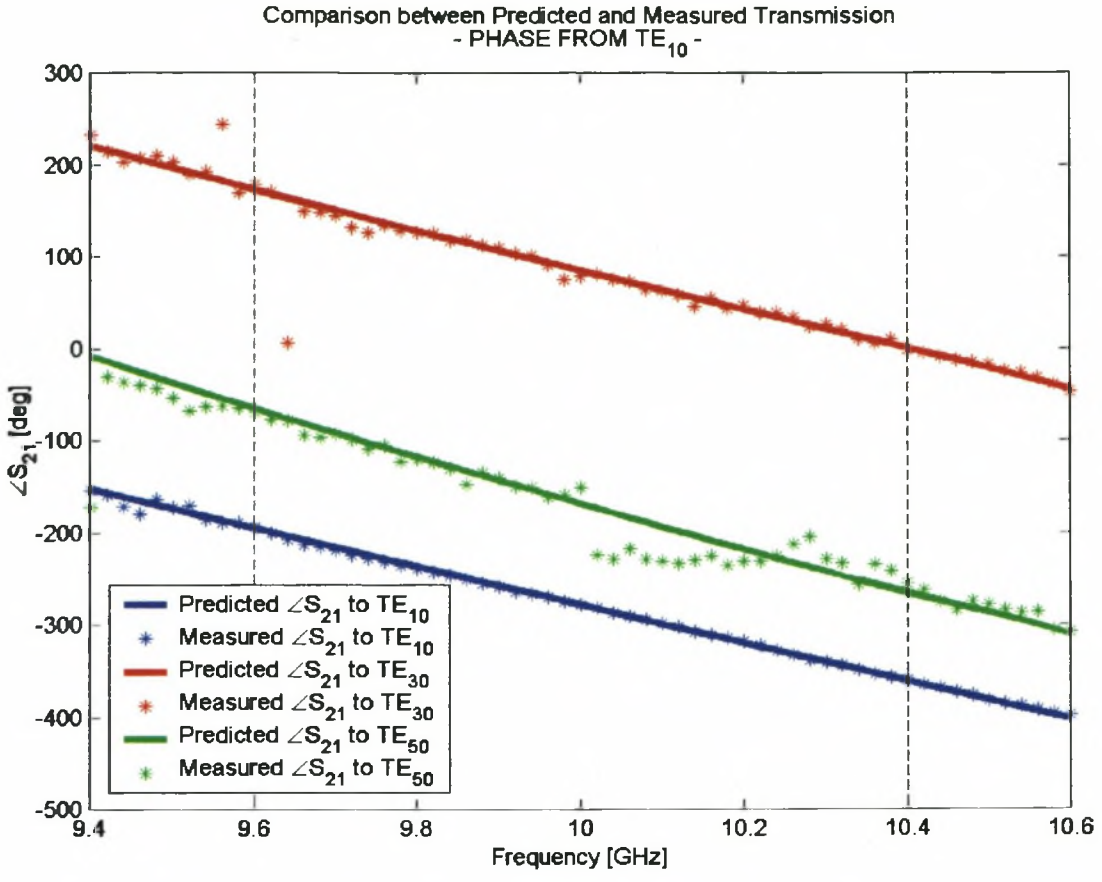


Figure 6.12 Predicted and measured transmission phase from TE_{10} mode to the TE_{10}, TE_{30} and TE_{50} modes at the output aperture

Although this horn exhibits no flaring in the E-plane, it is highly applicable to spatial power combining. Multiple copies of the horn could, for example, be stacked to build up an E-plane arrangement, or it could simply be flared in the E-plane with a smooth taper, since propagation speeds of the excited $TE_{(2n+1)0}$ modes are not dependent on waveguide height. The design compares favourably with existing hard horn designs, achieving a less than ±0.15dB variation of the aperture fields against the ±1dB variation obtained with the use of dielectric slabs, though over a smaller area of the aperture. If so desired, the design could have instead been optimised for maximal width and flatness of the aperture field. A great advantage of this design is the improved bandwidth and return loss which amounts to about 32dB maximum back-to-back over the 8% bandwidth. The aperture fields also achieve uniform flatness over the entire 8% band.

H-PLANE SPATIAL POWER COMBINING FEED (SURROGATE MODELS)

In order to illustrate the potential of the surrogate model analysis approach for functional blocks, proposed in section 3.6, the design above was repeated. This time the H-Plane Step functional block was made self-analysing, and used the interpolation-based surrogate model to determine its GSM. The surrogate model, which consisted of two physical variables; input and output step width, was trained at 324 points (18x18). The design time was significantly shorter than that of the full mode matching technique, but produced a horn of equal quality as the design above (according to the surrogate model analysis). To compare the effects of the surrogate model approximation, the design was re-analysed using mode matching. A diagram of the horn is shown in figure 6.13 along with its surrogate model and mode matching predicted return loss in figure 6.14. The modal weight magnitude ratios and phase offsets according to the surrogate model and the full mode matching model are shown in figure 6.15.

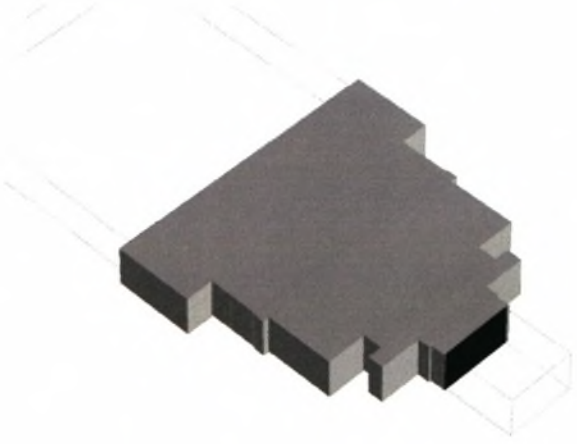


Figure 6.13 Schematic of Horn

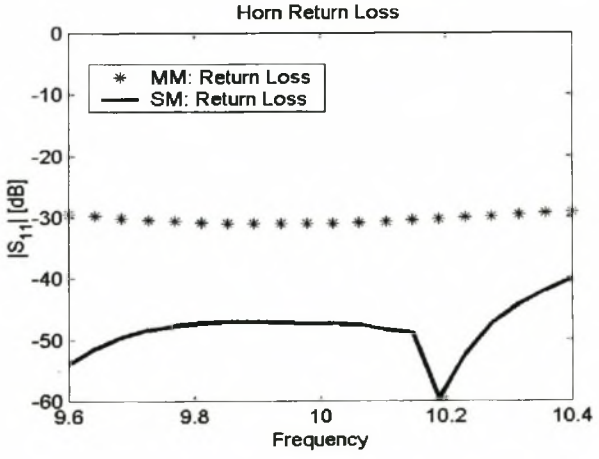


Figure 6.14 Return loss comparison

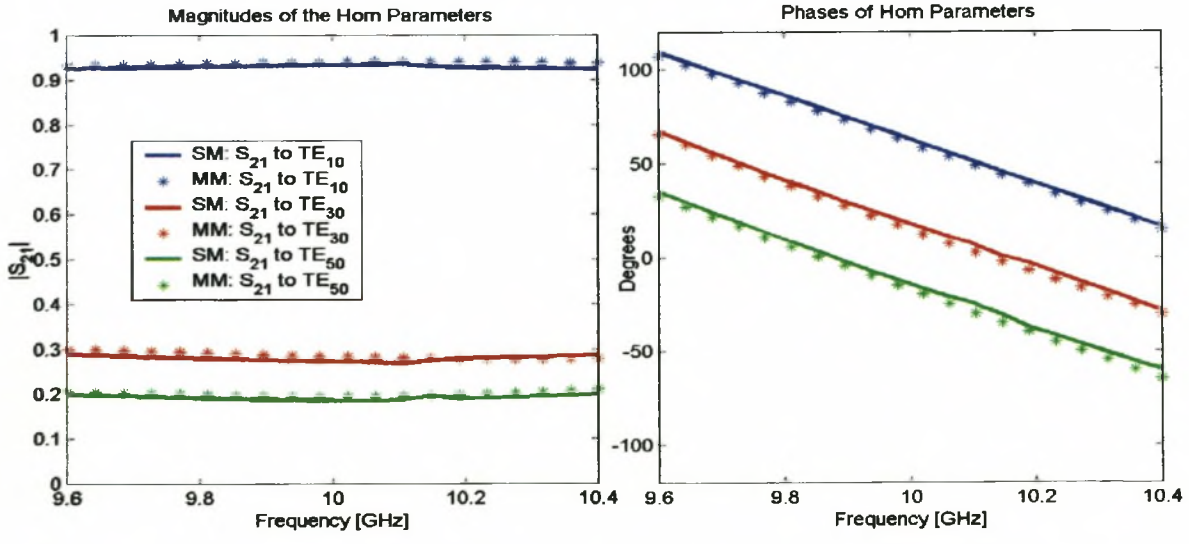


Figure 6.14 Comparison of Mode Matching (MM) and Surrogate Model (SM) analysis.

The use of an accurate surrogate model for this design resulted in a close correspondence between the mode matching analysis and the surrogate model analysis of the structure. Such small discrepancies are useful, as they allow the surrogate model-based design to be used as an excellent starting point for a final optimisation using full mode matching analysis. Clearly, lower accuracy surrogate models will result in worse starting points for this final stage of surrogate model assisted design.

It might be possible to make use of low accuracy surrogate models to develop rough designs at very high speed. As the design grows in complexity according to the strategies in chapter 2, the accuracy of the surrogate models could be gradually improved.

The application of extremely fast surrogate modelling in the context of the design strategies introduced here is potentially a powerful tool and is left as a subject of potential further research.

DOUBLE PLANE MULTIMODE HORN ANTENNA (MODE MATCHING)

The design problem described by Wriedt et al. calls for a double-plane flared horn antenna with maximal far field pattern symmetry [Wriedt 1989]. The proposed design approach makes use of an evolutionary algorithm to solve for the mode excitation substructure at the mouth of the horn. The same design problem is revisited here with the design strategies described in this chapter.

The horn due to Wriedt et al. achieves good pattern symmetry, as shown in figure 6.17, with a total physical length of about 20cm. Figure 6.19 shows the modal weight ratios achieved by the mode excitation substructure. At the design frequency (11.2 GHz) these modes combine to produce the far field patterns shown in figure 6.17. The pattern symmetry is, however, limited to a narrow band due to the strong change in modal weight ratios apparent further from the design frequency, as can be seen clearly for the TM_{12} mode in figure 6.19. A weakness of the horn, according to the authors, is a relatively high cross-polarization of –22dB, that appears to result from the choice of modal weights necessary to achieve pattern symmetry.

Since the phase relationships between the modal weights used by Wriedt et al. were not published, the optimal modal weights for this design problem had to be re-designed using the techniques described in section 6.3. In order to attempt to alleviate the cross-polarization problem, an additional criteria for the procedure was to find modal weights that provide the least possible excitation of x-oriented **E**-fields, whilst maintaining maximal pattern symmetry. The optimal modal weights that emerged, shown in figures 6.18 and 6.20, were somewhat different from those used by Wriedt et al., but result in cross-polarization of less than –30dB.

As with the examples above, the scanning technique was used in the design of this horn and the functional blocks were limited to symmetrical double-plane steps. In addition to controlling the excitation of the four modes used to create the radiation pattern, care also had to be taken to suppress other higher order modes propagating in the aperture, which could otherwise lead to high sidelobe levels.

Figure 6.15 shows the best result obtained after 12 scans. The horn is rather short at only 7.8 cm ($2.6\lambda_0$), and achieves good pattern symmetry with reasonably low side lobes, as shown in figure 6.16. The return loss is shown in figure 6.21 and the achieved and optimal modal

weight magnitude ratios and phase relationships are shown in figure 6.18 and 6.20 respectively. Figure 6.22a) shows the complete far field radiation pattern and figure 6.22b) shows the cross-polarized radiation pattern.

This design consumed considerable computational resources, taking about six days to complete only 12 scans on a 1.4 GHz machine. The used of surrogate model approximations of the double plane steps used as functional blocks might have alleviated this, allowing more scans and a better design solution.

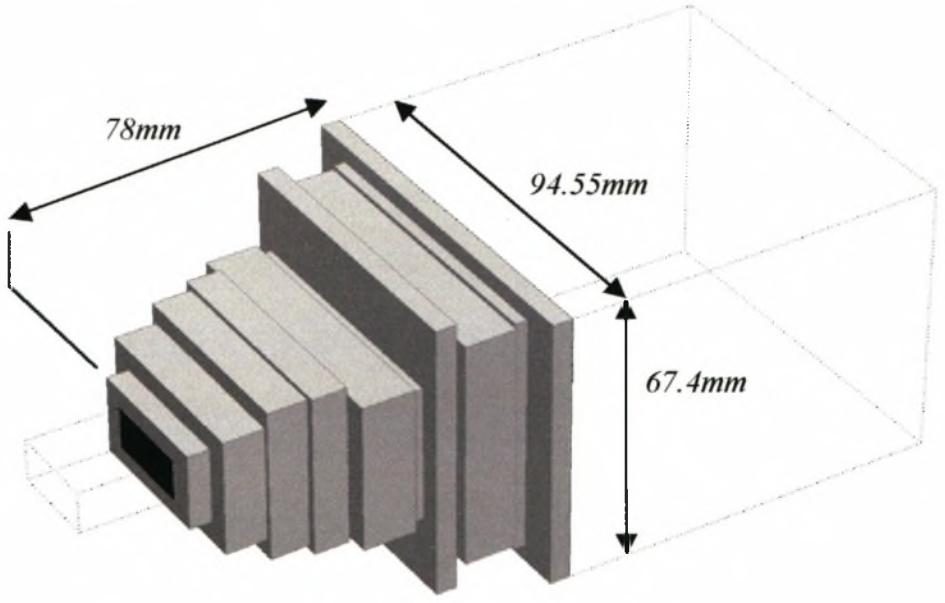


Figure 6.15 Diagram of the designed double-plane horn

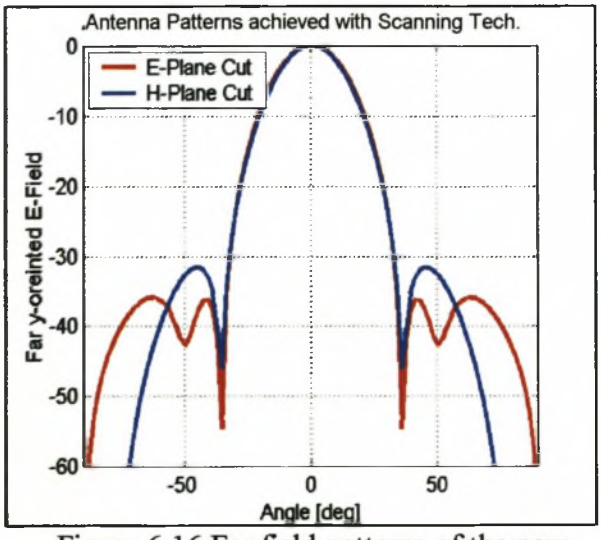


Figure 6.16 Far field patterns of the new horn antenna

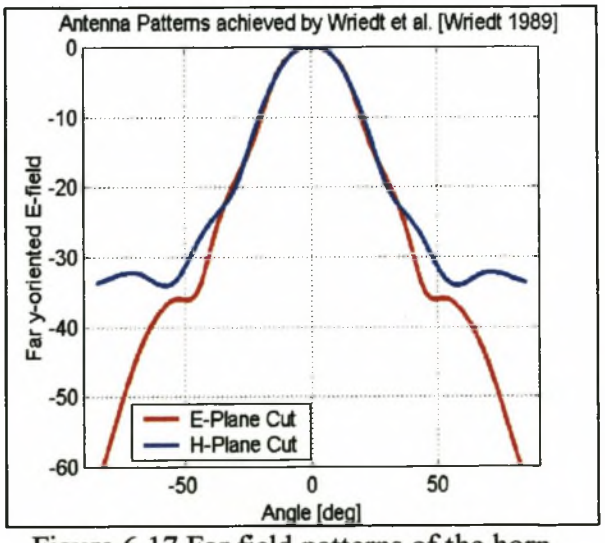


Figure 6.17 Far field patterns of the horn due to Wriedt et al. [Wriedt 1989]

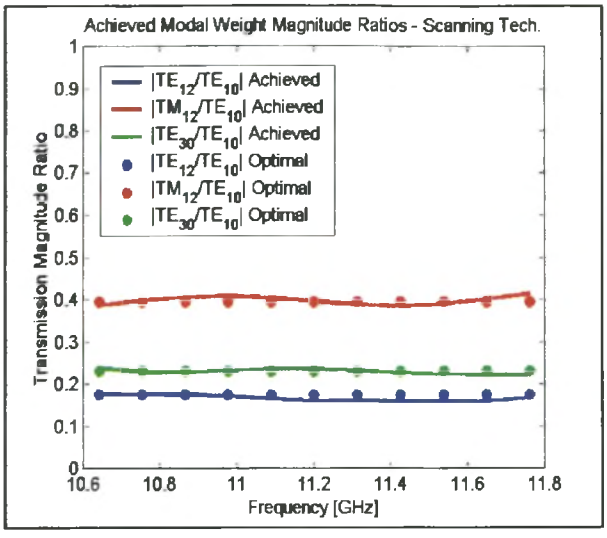


Figure 6.18 Designed magnitude ratios

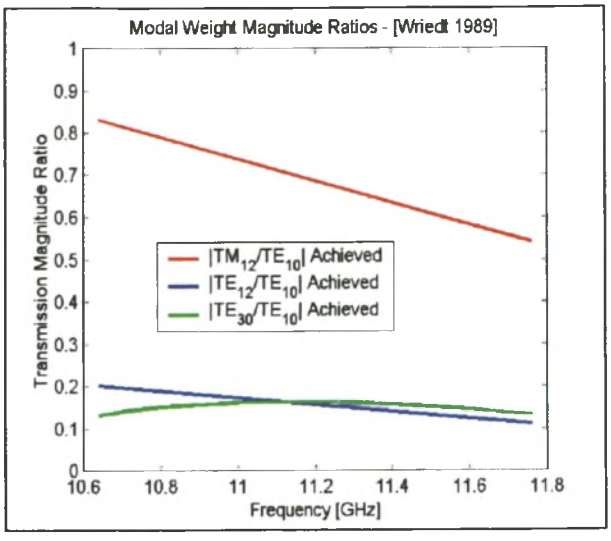


Figure 6.19 Mag. ratios used by Wriedt et al.

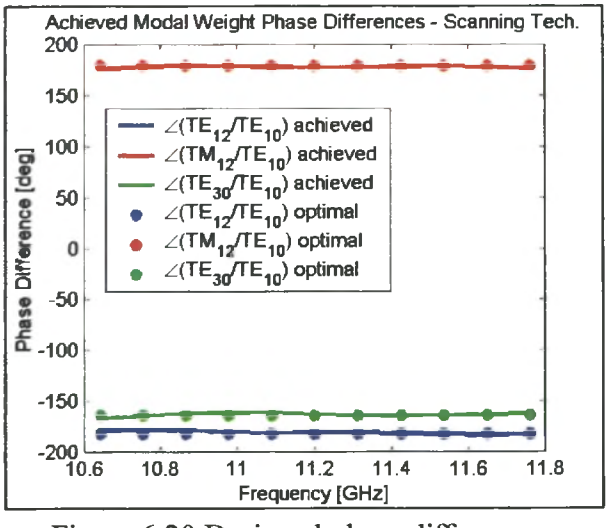


Figure 6.20 Designed phase differences

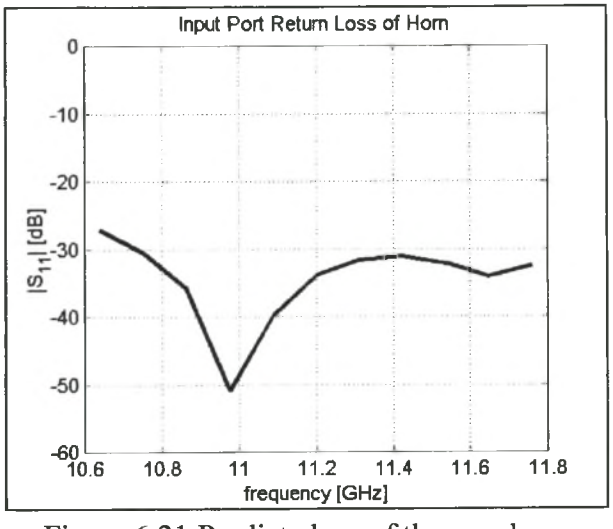


Figure 6.21 Predicted s_{11} of the new horn

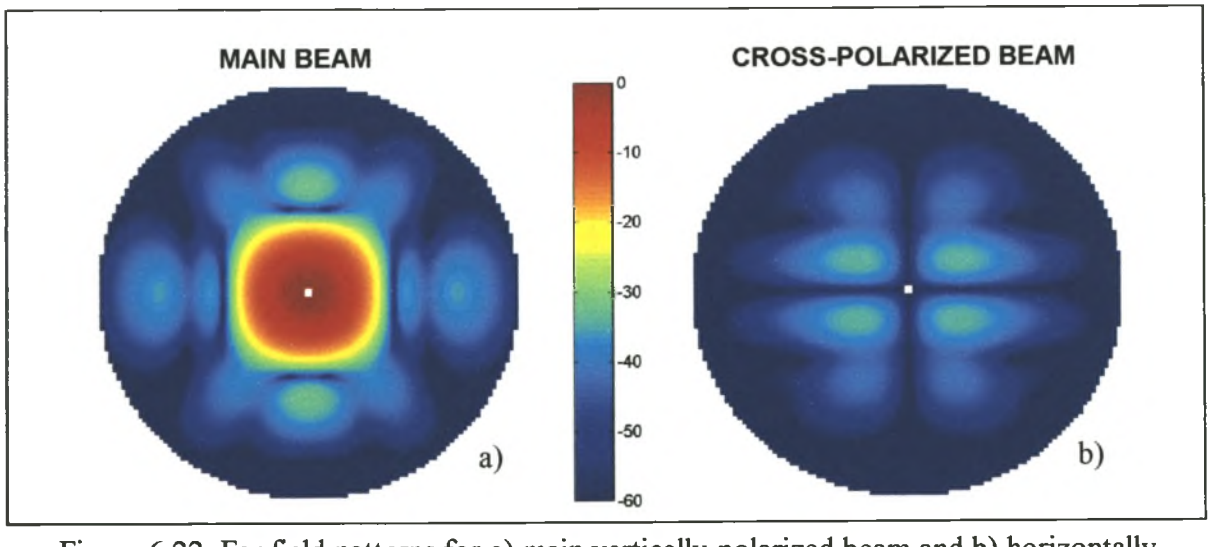


Figure 6.22. Far field patterns for a) main vertically-polarized beam and b) horizontally polarized beam

6.7 Conclusion

This chapter has introduced a new design technique for multimode horns. The design starts by determining, through optimization, the applicable modal weights to yield a desired aperture field pattern. It then goes on to use a growth-based scanning strategy to guide the gradual development of the structure through the successive addition of functional blocks. Although this investigation has experimented only with symmetrical waveguide steps, the use of other functional blocks, such as irises and stubs might provide increased performance advantages at the cost of a more time consuming and involved design process.

Three design examples were presented which illustrate the potential of the technique to provide improved designs for both spatial power combining applications and in pattern shaped antennas and feeds. Surrogate model assisted analysis was employed in one of the designs to illustrate its application to fast designs.

Areas for further research are the potential application of surrogate models in a space-mapping oriented design technique, and the application of the proposed design technique to horns and feeds with both multiple input and output modes, such as would be used in a tracking radar. Such a design might have to incorporate asymmetric waveguide steps or other types of functional block to accomplish successful design.

Chapter 7

Evaluation, Outlook and Conclusions

7.1 General Evaluation of the CAD Techniques

The CAD technique proposed in this thesis was successfully applied to solve general design problems associated with two practical multimode systems; multimode bandstop filters and multimode horns. As yet, the strategies proposed in this investigation appear to be the most viable design techniques for these particular examples, with both implementations representing original contributions to their respective fields.

The reliability and efficiency of the various strategies proposed here have already been evaluated in conjunction with the specific examples to which they are applied, but these aspects deserve more general comment in this discussion.

7.1.1 Quality of the design solutions

The preferred outcome of any design technique is a high quality solution. Such a solution should not only fulfil electrical demands, but should also have good structural qualities, such as small physical dimensions and a simple structure facilitating easy fabrication. The design strategies presented here are oriented at pursuing 'globally' optimal design. Whether this is generally accomplished through the use of growth-based techniques is difficult to ascertain since the examples pursued in this investigation have no known optimal solutions.

When compared with existing alternative designs for these problems, the emergent solutions showed a number of favourable electrical characteristics. However, it is fair to say that such improvements came at the cost of a rather complex structure, consisting of typically large amounts of arbitrarily placed discontinuities of varying type in a single cascade structure. The trade-off between this increased complexity and the improvement of existing electrical standards depends largely on the application. In situations where there simply is no other

viable design approach, the increased physical complexity of the solution structures may well not be an issue. In some cases, it might occur that the solution structures developed using this approach could replace even more unwieldy alternatives, such as tapered structures.

7.1.2 Reliability of the CAD Strategies

The strategies developed here were generally found to be reliable. This is most probably due to the growth-based approach pursued, which tended to avoid the traps of local minima. Although the stochastic nature of the design strategies ensured minor differences in outcomes of repeated runs of various design problems, the successful outcomes were typically observed to exhibit global structural similarities. The growth based approach, cannot, however, ensure 100% percent reliability and some outcomes were unacceptable. The origin of these bad solutions might well be as a result of an inferior optimization tool. It was observed that when the move was made from using a least squares optimization to a hybrid least squares, steepest descent algorithm, the proportion of inferior outcomes dropped dramatically in the design examples presented in this thesis.

When applied to known problems, such as single mode filter design, the strategies were found to generally reproduce topologies encountered as a result of synthesis (with the expenditure of a deal more time), but it was observed that very narrow bandwidth designs suffered from lower quality solutions and sometimes failed to produce acceptable results. This is possibly due to the extremely narrow range of physical dimensions that must be pursued in such a design. Results could generally be improved by sacrificing speed of design for a more rigorous search of the solution space through the incorporating of the some of the enhancements discussed in chapter 2, by forcing more scans in the scanning strategy, or by using a larger population in the evolutionary strategy.

In light of the difficulties with single mode, highly frequency selective design, it is unlikely that similarly natured multimode problems would be any easier to solve with the techniques proposed here. Accordingly, these techniques cannot guarantee reliability in highly frequency selective problems. It may, however, be possible to overcome these problems, through the formulation of more advanced error functions.

7.1.3 Application of the design approach to other problems

The success of the design techniques in dealing with the considerably different types of multimode problems detailed in chapters 5 and 6 suggests that these design techniques might be suitable for a wide range of other multimode problems, including the design of mode converters and overmoded bandpass filters.

The flexibility of the approach is due to the minimal information that is required regarding the desired electrical characteristics of the devices and the use of the functional block based-description of the designed structure. These properties make it attractive to other design problems which also do not conform to standard approaches. An example of such application is the design of a single mode frequency discriminator, which provides a perfectly linear frequency selectivity over frequency. The scanning technique successfully produced a viable passive structure that accomplishes this task.

7.2 An Outlook to Further Research

This work has opened a number of opportunities for further research. These are not limited solely to the applications of multimode design procedures, but also to enhancements of existing measurement procedures and numerical analysis techniques. The specific areas of potential further research that have been prompted by this investigation are:

- **Further investigation of the mode matching equivalent circuit and its potential applications to modelling and design**. The circuit provides an excellent physical insight into the mode matching operation that might well be exploited at a design level.

- **The extension of the reduced mode matching approach so as to avoid needless computations of the W-matrix entries**. The reduced mode matching formulation in its current implementation requires complete evaluation of the W-matrix before reduction can take place. A modification of the technique is possible that does not require this. This is important for applications which make use of numerical integration to determine the matrix entries, such as in circular waveguide boundaries.

- **The more complete incorporation and understanding of surrogate models as applied to the design procedure developed in this investigation and the potential use of a space mapping based approach for the gradual improvement of the surrogate model as design progresses.** The basic surrogate model implementation has shown great promise as a potential technique of reducing the computational effort of multimode design. However, its properties and potential application need to be further understood.

- **The extension of multimode bandstop filter design to incorporate phase compensation functional blocks to provide improved electrical performance.** The very simplest implementation of multimode bandstop filter design was investigated in this research. The growth-based CAD strategies presented here offer the potential of greatly improving these simple design through the inclusion of alternative functional blocks. This could provide advantages in terms of performance and structural properties.

- **The development of a practical spatial power combiner incorporating multimode horns feeds.** The examples implanted in chapter 6 did not take into account the effect of such disturbances as the active array and the back-to-back arrangement of the horns. The examples served to show what was possible with the techniques available. The strong degree of control the design techniques have of managing modal weights and phases could be utilized for the purposes of compensating mismatches to the active array and surrounding structures. There is also the possibility of the incorporation of frequency selectivity directly in the horn design, that might, for example be geared to distribute power at different locations in the array for different excitations.

- **Application of the design techniques to monopulse radar feeds.** A strength of the proposed design techniques is the ability to deal with multiport problems. In this investigation, this was applied to multiple electrical ports in single apertures at the input and output of the device. The application of similar techniques might be possible to problem in multiple physical port problems, such as encountered in monopulse feed problems. This would require a more general functional block based formulation to be developed, but the underlying growth based strategies are still viable.

- **Extension of the CAD strategies proposed in chapter 2.** A number of additional enhancements to these techniques are possible, incorporating Knowledge-Aided Design (KAD) as well as more systemic searches, through the implementation of the generalized tree search scanning strategy, and more reliable optimization algorithms. The potential for space mapping optimization to reduce design times involving the scanning strategy should also be investigated.

- **Internal Signal Matching.** Of the many design approaches that had to be discarded for not being sufficiently applicable to a large variety of specifications, this was by far the most interesting. Where normal scanning decides on the inclusion of a block based on its ability to effect only the input/output signal relationship of the structure, *internal signal matching* selects a block based on how well it matches *all* the internal signals of the structure – or internal field distributions to those necessary for successful output. This is done by exciting the structure at the output port by the desired output, and at the input port by the desired input, and seeing how well a particular functional block inserted at an arbitrary location causes these signals to match through the structure. Tests with this scheme showed excellent convergence to viable solution structures that surpassed those obtained with the scanning technique. A variation on the technique designs an insertion structure to best match the signals, instead of simply placing a functional block. The general technique is applicable to multiport structures, which is one requirement of the investigation, but it, unfortunately, relies upon a full output specification, which cannot be guaranteed.

7.3 Conclusion

A new method for the design of multimode devices in overmoded waveguide is presented in this thesis. The design makes use of the principle of growth-based design and incorporates a functional block based description of the design problem to facilitate effective and flexible design. Two general design strategies are implemented and tested. The first is a scanning-based technique and the second uses an evolutionary strategy to gradually introduce complexity into the design solution as the design progresses.

For the design to be viable on a computational level, novel analysis techniques were developed. The first uses a reduced mode matching analysis, incorporating intelligent selection of appropriate modes to reduce computational effort. The second models the functional blocks with multiple-output oriented interpolation-based surrogate model that is trained prior to design with the mode matching technique.

The challenges which make design of multimode devices difficult also hold for the measurement of these devices. An original method of efficiently measuring the full scattering parameters of a multimode devices was developed for use with a standard 2-port network analyzer. The technique was used to measure successfully a number of overmoded devices implemented with the design techniques described in this thesis.

The growth-based design strategies were successfully applied to two practical problems of multimode design, resulting in original methods of design for both types of problems. The first problem was the design of multimode bandstop filters for applications as waveguide chokes. The solution to this problem identified fixed resonant functional blocks that were cascaded with the aid of the scanning and evolutionary strategies. The second problem was the design of multimode horns, typically applicable to antennas and antenna feeds, as well as in the developing field of spatial power combining. The solution strategy placed waveguide step functional blocks using the scanning strategy.

The design strategies developed are by no means limited to the problems discussed in this thesis. It is hoped that future research will investigate the solution of other multimode design problems using the concepts and tools detailed here.

References

[Accatino 1994]: L. Accatino and G. Bertin, "Design of Coupling Irises Between Circular Cavities by Modal Analysis," *IEEE Trans. Microwave Theory Tech.*, vol. 42, no. 7, pp. 1307-1313, July, 1994.

[Alessandri 1995]: F. Alessandrim M. Dionigi and R. Sorrentino, "A Fullwave CAD Tool for Waveguide Components Using a High Speed Direct Optimizer," *IEEE Trans. Microwave Theory Tech.*, vol.43, no. 9, pp. 2046-2052, Sept. 1995.

[Alessandri 1988]: F. Alessandri, G. Bartolucci and R. Sorrentino, "Admittance Matrix Formulation of Waveguide Discontinuity Problems: Computer-Aided Design of Branch Guide Couplers," *IEEE Trans. Microwave Theory Tech.*, vol. 36, no. 2, pp. 394-403, Feb 1988.

[Alessandri 1992]: F. Alessandri, M. Mongiardo, R. Sorrentino, "Computer-Aided Design of Beam Forming Networks for Modern Satellite Antennas," *IEEE Trans. Microwave Theory Tech.*, vol. 40, no. 6, pp. 1117-1127, June 1992.

[Alessandri 1997]: F. Alessandri, M. Dionigi, R. Sorrentino and L. Tarricone, "Rigorous and Efficient Fabrication-Oriented CAD and Optimization of Complex Waveguide Networks," *IEEE Trans. Microwave Theory Tech.*, vol. 45, no. 12, pp. 2366-2374, Dec. 1997.

[Ali 1998]: A. Ali, S. Ortiz, T. Ivanov, and A. Mortazawi, "Analysis and measurement of hard horn feeds for the excitation of quasi-optical amplifiers," *IEEE Int. Microwave Symp. Dig.*, pp 1469-1472, June 1998.

[Arndt 1997]: F. Arndt, R. Beyer, J.M. Reiter, T. Sieverding and T. Wolf, "Automated Design of Waveguide Components Using Hybrid Mode-Matching/Numerical EM Building-Blocks in Optimization-Oriented CAD Frameworks – State-of-the-Art and Recent Advances," *IEEE Trans. Microwave Theory Tech.*, vol. 45, no. 5, pp. 747-759, May 1997.

[Atia 1971]: A.E. Atia and A.E. Williams, "New Types of Waveguide Bandpass Filters for Satellite Transponders," *COMSAT Technical Review*, vol. 1, no. 1, pp. 21-42, Fall, 1971.

[Bäck 1997]: T. Back, U. Hammel and H. Schweifel, "Evolutionary Computation: Comments on the History and Current State," *IEEE Trans. Evol. Comp.*, vol. 1, no. 1, pp. 3-17, April 1997.

[Bakr 2000]: M.H. Bakr, J.W. Bandler, K. Madsen, J.E. Rayas-Sanchez and J. Sondergaard, "Space-mapping optimization of microwave circuits exploiting surrogate models," *IEEE Trans. Microwave Theory Tech.*, vol. 48 no. 12, pp. 2297-2306 Dec. 2000.

[Bandler 1995]: J.W. Bandler, R.M. Biernacki, S.H. Chen, R.H. Hemmers and K. Madsen, "Electromagnetic optimization exploiting aggressive space mapping," *IEEE Trans. Microwave Theory Tech.*, vol. 43, pp.2874-2882, 1995.

[Barrow 1939]: W.L. Barrow and L.J. Chu, "Theory of the electromagnetic horn," *Proc. Inst. Radio Eng.*, no. 27, pp 51-64, 1939.

[Beck 1955]: A.C. Beck, "Measurement Technique for Multimode Waveguides," *IRE Trans. Microwave Theory Tech.*, pp. 35-41, April, 1955.

[Buckley 1990]: M. J. Buckley and R. J. Vernon, "Compact Quasi-Periodic and Aperiodic TE_{0n} Mode Converters in Overmoded Circular Waveguides for Use with Gyrotrons," *IEEE Trans. Microwave Theory Tech.*, vol. 38, no. 6, pp 712-721, June, 1990.

[Bunger 1999]: R. Bunger, R. Beyer and F. Arndt, "Rigorous Combined Mode-Matching Integral Equation Analysis of Horn Antennas with Arbitrary Cross Section," *IEEE Trans. Antennas Propagat.*, vol. 47, no. 11, pp. 1641-1648, Nov. 1999.

[Butterweck 1968]: H. Butterweck, "Mode Filters for Oversized Rectangular Waveguides," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-16, no. 5, pp. 274-281, May 1968.

[Carmel 1984]: Y. Carmel, K. R. Chu, M. E. Read, V. L. Granatstein, G. Faillon, P. Boulanger, E. Kammerer, and G. Mourier, "A Technique to Identify Electromagnetic Modes in Oversize Waveguides," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-32, pp. 1493-1495. Nov. 1984.

[Collin 1985]: R.E. Collin, Antennas and Radiowave Propagation, McGraw-Hill, Singapore, pp. 164-171, 1985.

[CMPM 1994]: Committee on Microwave Processing of Materials: An Emerging industrial technology, National Materials Advisory Board, Commission on Engineering and Technical Systems and National Research Council, Microwave Processing of Materials, (Publication NMAB-473), National Academy Press, Washington D.C. pp. 42-43, 1994.
(on WWW at : http://books.nap.edu/books/0309050278/html/42.html#pagetop)

[Clarricoats 1981]: P.J.B. Clarricoats and R.D. Elliot, "Multimode corrugated waveguide feed for monopulse radar," *IEE Proc.*, vol. 128, Pt. H, no. 2, pp 102-110, April 1981.

[Doane 1988]: J.L. Doane, "Low-Loss Twists in Oversize Rectangular Waveguide," *IEEE Trans. Microwave Theory Tech.*, vol. 36, no. 6, pp. 1033-1042, June 1988.

[Eisenhart 1971]: R.L Eisenhart and P.J. Khan, "Theoretical and Experimental Analysis of a Waveguide Mounting Structure," *IEEE Trans. on Microwave Theory Tech.*, vol. MTT-19, no. 8, pp. 706-719, August 1971.

[Eleftheraides 1994]: G.E. Eleftheraides, A.S. Omar, L.P.B. Katehi and G.M. Rebeiz, "Some Important Properties of Waveguide Junction Generalized Scattering Matrices in the Context of the Mode-Matching Technique," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-42, no. 10, pp. 1896-1903, October 1994.

[Felsen 1959]: L. B. Felsen, W. K. Khan and L. Levey, "Measurement of Two-Mode Discontinuities in a Multimode Waveguide by a Resonance Technique," *IRE Trans. Microwave Theory Tech.*, pp. 102-110, Jan., 1959.

[Granet 2000]: C. Granet, T.S. Bird and G.L. James, "Compact Multimode Horn with Low Sidelobes for Global Earth Coverage," *IEEE Trans. Antennas Propagat.*, vol. 48, no. 7, pp. 1125-1133, July 2000.

[Guglielmi 1994]: M. Guglielmi, "Simple CAD Procedure for Microwave Filters and Multiplexers," *IEEE Trans. Microwave Theory Tech.*, vol. 42, no. 7, pp. 1347-1352, July 1994.

[Gupta 1998]: K.C. Gupta, "Emerging trends in millimetre-wave CAD," *IEEE Trans. Microwave Theory Tech.*, vol. 46, no. 6, pp 747-755, June 1998.

[Hannan 1961a]: P.W. Hannan, "Optimum Feeds for All Three Modes of a Monopulse Antenna I: Theory," *IRE Trans. Antennas Propagat.*, pp 444-454, Sept. 1961.

[Hannan 1961b]: P.W. Hannan, "Optimum Feeds for All Three Modes of a Monopulse Antenna II: Practice," *IRE Trans. Antennas Propagat.*, pp 454-461, Sept. 1961.

[Harvey 2000]: J. Harvey, E.R. Brown, D.B. Rutledge and R.A. York, "Spatial Power Combining for High-Power Transmitters," *IEEE Microwave Magazine*, pp. 48-59, Dec. 2000.

[Harrington 1961]: R.F. Harrington, Time Harmonic Electromagnetic Fields, McGraw-Hill, New York, 1961, p143

[Hsu 2000]: Hsu, Jui-Pang, T. Hiraoka and H. Honma, "Equivalent network for rectangular-waveguide H-plane step discontinuity – Multi-transmission line and Multi-port ideal transformer – ," *IEEE MTT-S Intl. Symp. Digest*, vol. 2., pp. 1069-1072, Boston, 11-16 June 2000.

[Ivanov 1996]: T. Ivanov and A. Mortazawi, "A Two-Stage Spatial Amplifier woth Hard Horn Feeds," *IEEE Microwave and Guided Wave Letters*, vol. 6, no. 2, pp. 88-90, Feb. 1996.

[Kelley 1999]: C.T. Kelley, Iterative Methods for Optimization (in applied mathematics), Society for Industrial and Applied Mathematics (SIAM), 1999.

[Koza 1997]: J.R. Koza, F.H. Bennett, D. Andre, M.A. Keane and F. Dunlap, "Automated Synthesis of Analog Electrical Circuits by Means of Genetic Programming," *IEEE Trans. Evol. Comp.*, vol. 1, no. 2, pp. 109-128, July 1997.

[Kühn 1984]: E. Kühn and B.K. Watson, "Rectangular Corrugated Horns – Analysis, Design and Evaluation," *Proc. 14th European Microwave Conf.*, pp. 221-227, Sept. 1984.

[Lehmensiek 2001]: R. Lehmensiek, P. Meyer, "Creating accurate multivariate rational interpolation models of microwave circuits by using efficient adaptive sampling to minimize the number of computational electromagnetic analyses," *IEEE Trans. Microwave Theory Tech.*, vol. 49, no. 8, pp. 1419-1430, Aug. 2001.

[Lee 1988]: K.M. Lee and R-S Chu, "Design and Analysis of a Multimode Feed Horn for a Monopulse Feed," *IEEE Trans. Antennas Propagat.*, vol. 36, no. 2, pp 171-181, Feb. 1988.

[Leroy 1983]: M. Leroy, "On the convergence of numerical results in modal analysis," *IEEE Trans. Antennas Propagat.*, vol. AP-31, pp. 655-659, July 1983.

[Levinson 1966]: D.S. Levinson and I. Rubinstein, "A Technique for Measuring Individual Modes Propagating in Overmoded Waveguide," *IEEE Trans. Microwave Theory Tech.*, vol. 14, pp. 310-322, July 1966.

[Lier 1986]: E. Lier, "Cross Polarization from Dual Mode Antennas," *IEEE Trans. Antennas Propagat.*, vol. AP-34, no. 1, pp 106-110, Jan 1986.

[Love 1962]: A.W. Love, "The Diagonal Horn Antenna," *Microwave Journal,* vol. V, no. 3, pp 117-122, March 1962.

[Lunéville 1998]: E. Lunéville, J Kreig and E. Giguet, "An Original Approach to Mode Converter Optimum Design," *IEEE Trans. Microwave Theory Tech.,* vol. 46, no. 1, pp 1-9, Jan., 1998.

[Marcuvitz 1951]: N. Marcuvitz, The Waveguide Handbook, McGraw-Hill, 1951

[Metaxas 1983]: A.C. Metaxas and R.J. Meredith, Industrial microwave heating, Peregrinus London, 1983.

[Meyer 1995]: P. Meyer, A Combined Mode-Matching and Method of Lines Procedure for the Analysis of Planar Microwave Circuits, *Ph.D. Dissertation,* University of Stellenbosch, 1995.

[Olver 1994]: A.D. Olver, P.J.B. Clarricoats, A.A. Kishk and L. Shafai, Microwave Horns and Feeds, IEEE Press, New York, 1994, pp. 229-255.

[Omar 1985]: A.S. Omar and K. Schunemann, "Tarnsmission Matrix Representation of Finline Discontinuities," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-34, no.9, pp. 765-770, Sept. 1985.

[Ortiz 2000]: S. Ortiz, J. Hubert, L. Mirth, E. Schlecht, A. Mortazawi, "A 25 watt and 50 Watt Ka-band quasi-optical amplifier," *Proc. IEEE MTT-S Int. Microwave Symp.*, Boston, MA., June 11-16, 2000, vol. 2, pp. 797-800.

[de Pádua 2000]: R. de Padua, L.R.A.X De Menezes, "Direct synthesis of microwave filters using inverse scattering TLM (transmission line matrix) method," *Proc. IEEE MTT-S Int. Microwave Symp.*, Boston, MA., June 11-16, 2000, vol. 1, pp. 379-382

[Polak 1997]: E. Polak, Optimization Algorithms and Consistent Approximations, Springer-Verlag, New York, 1997.

[Potter 1963]: P.D. Potter, "A New Horn Antenna with Suppressed Sidelobes and Equal Beamwidths," *Microwave Journal,* no. 6, pp 71-78, June 1963.

[Price 1959]: V. G. Price, "Measurement of Harmonic Power Generated by Microwave Transmitters," *IRE Trans. Microwave Theory Tech.,* pp. 116-120, Jan., 1959.

[Ramo 1993a] S. Ramo, J. R. Whinnery and T. Van Duzer, Fields and Waves in Communication Electronics, (3rd edition), John Wiley and Sons, New York, 1993, pp. 396-397.

[Ramo 1993b] S. Ramo, J. R. Whinnery and T. Van Duzer, Fields and Waves in Communication Electronics, (3rd edition), John Wiley and Sons, New York, 1993, pp. 417-422.

[Ratner 1996]: I. Ratner, H.O. Ali, E.M. Petriu and G. Eatherley, "Neural Network modelling of electromagnetic field problems," *in Proc. Internat. Workshop on Neural Networks for Identification, Control, Robotics, and Signal/Image Processing*, pp. 387-391, 1996

[Reed 1956]: J. Reed and G.J. Wheeler, "A Method of Analysis of Symmetrical Four-Port Networks," *IRE Trans. on Microwave Theory Tech.*, pp. 246-252, October 1956.

[Ren 1974]: C. Ren and H. Wang, "A Class of Wavguide Filters for Over-Moded Applications," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-22, no. 12, pp. 1202-1209, Dec. 1974.

[Rosenberg 2000]: U. Rosenberg and M. Schneider, "High-Performance Transitions for Overmoded Opertaion of Elliptical Waveguides," *IEEE Trans. Microwave Theory Tech.*, vol. 48, no. 10, pp. 1749-1755, Oct. 2000.

[Roberts 1995]: P.P. Roberts and G.E. Town, "Design of Microwave Filters by Inverse Scattering," *IEEE Trans. Microwave Theory Tech.*, vol. 43, no. 4, pp. 739-742, April 1995.

[Sequinot 1998]: C. Seguinot, P. Kennis, J. Legier, F. Huret, E. Paleczny, and L. Hayden, "Multimode TRL – A New Concept in Microwave Measurements: Theory and Experimental Verification," *IEEE Trans. Microwave Theory Tech.*, vol. 46, pp. 536-542, May, 1998.

[Sharp 1961]: E. D. Sharp and E. M. T. Jones, "A Sampling Technique for the Measurement of Multimode Harmonic Waveguide Power," *IRE Trans. Microwave Theory Tech.*, pp. 73-82, Jan., 1961.

[Shih 1989]: Y.C. Shih, Numerical Techniques for Microwave and Millimetre-Wave Passive Structures, (ed. T. Itoh), John Wiley and Sons, New York, 1989, Chapter 9.

[Sorrentino 1991]: R. Sorrentino, M. Mongiardo, F Alessandri and G. Schiavon, "Investigation of the numerical properties of the mode-matching technique," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 4, no. 1, pt. 2, pp. 19-43, March 1991.

[Soto 2000]: P. Soto, V.E. Boria, J.M. Catala-Civera, N. Chouaib, M. Guglielmi and B. Gimeno, "Analysis, Design, and Experimental Verification of Microwave Filters for Safety Issues in Open-Ended Waveguide Systems," *IEEE Trans. Microwave Theory Tech.*, vol. 48, no. 11, pp. 2133, 2140, Nov. 2000.

[Tang 1996]: K.S. Tang, K.F. Man S. Kwong and Q. He, "Genetic Algorithms and their Applications," *IEEE Signal Processing Magazine*, pp. 22-37, Nov. 1996.

[Tascone 2000]: R. Tascone, P. Savi, D. Trinchero and R. Orta, "Scattering Matrix Approach for the Design of Microwave Filters," *IEEE Trans. Microwave Theory Tech.*, vol. 48, no. 3, pp. 423-429, March 2000.

[Taub 1962]: J. J. Taub and J. Goldberg, "A New Technique for Multimode Power Measurement," *IRE Trans. Microwave Theory Tech.*, pp. 64-69, 1962

[ul Haq 1995]: T. ul Haq, K.J. Webb, and N.C. Gallagher, "Scattering Optimization Method for the Design of Compact Mode Converters for Waveguides," *IEEE Trans. Microwave Theory Tech.*, vol. 43, no. 3, pp. 559-565, March 1995.

[Vale 1999]: C.A.W. Vale and P. Meyer, "Designing High-Performance Finline Tapers with Vector-based Optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 47, no. 12, pp. 2467-2462, Dec 1999

[Vale 2000a]: C.A.W. Vale and P. Meyer, "Modal Transforms in Rectangular Waveguide," *IEE Electronics Letters,* vol. 36 No. 16, pp. 1350-1352, August 2000.

[Vale 2000b]: C.A.W. Vale and P. Meyer, "A Design Technique for Co-existing Multimode Waveguide Bandstop Filters", *Proc. IEEE MTT-S Int. Microwave Symp.*, Boston, MA., June 11-16, 2000 pp. 1189-1192

[Vale 2000c]: C.A.W. Vale, P. Meyer and K. D. Palmer, "A Design Procedure for Bandstop Filters in Waveguides supporting Multiple Propagating Modes," *IEEE Trans. Microwave Theory Tech.*, vol. 48, no. 12, pp. 2496-2503, Dec., 2000.

[Vale 2000d]: C.A.W Vale and P.Meyer, "Waveguide Chokes for Microwave Heating Applications," *in Proc. SATCAM 2000*, Cape Town, South Africa, August 2000.

[Vale 2001]: C.A.W Vale and P.Meyer "Automated Intelligent Mode Selection for Fast Mode Matching Analysis of Waveguide Discontinuities," *IEEE MTT-S Intl. Symp. Digest,* Phoenix, 20-25 May 2001.

[Watson 1997]: P.M. Watson and K.C. Gupta, "Design and optimization of CPW circuits using EM-ANN models for CPW components" *IEEE Trans. Microwave Theory Tech.*, vol. 45, no. 12, pt 2., pp. 2515-2523, Dec. 1997.

[Weisshaar 1996]: A. Weisshaar, M. Mongiardo, A. Tripathi and V. K. Tripathi, "CAD-Oriented Fullwave Equivalent Circuit Models for Waveguide Components and Circuits," *IEEE Trans. Microwave Theory Tech.*, vol. 44, no. 12, pp 2564-2569, Dec. 1996.

[Wexler 1967]: A. Wexler, "Solution of Waveguide Discontinuities by Modal Analysis," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-15, pp. 509-517, 1967.

[Wriedt 1989]: T. Wriedt, K. Wolff, F. Arndt and U. Tucholke, "Rigorous Hybrid Field Theoretic Design of Stepped Rectangular Waveguide Mode Converters Including the Horn Transitions into Half-Space," *IEEE Trans. Antennas Propagat.*, vol. 37, no. 6, pp. 780-790, June 1989.

Appendix 1

General Purpose Optimization Tool

A1.1 Introduction

This investigation makes extensive use of optimization, not only in the design algorithms developed in chapter 2, but also in a number of secondary problems which involve the use of iterative solvers, such as measurement and the development of design goals. An efficient hybrid least squares, steepest descent algorithm has been developed that provides reliable convergence for most of these applications. The aim of the optimizer is not necessarily to provide global minima seeking properties, but rather to allow the reasonably efficient and reliable location of a nearby local minima, as is typical of most gradient-based optimization algorithms. The method proposed here is not essential to the design techniques developed in this dissertation, but has been found to confer an increased level of reliability to the strategies that use it.

A1.2 Evaluation of various optimization techniques

In order to evaluate various optimization techniques available, a standard problem, typical of the type encountered in this investigation, was formulated. A stepped horn structure, like that shown in figure A1.1 was created and analysed with the techniques discussed in chapter 3. This structure was then perturbed by varying its dimensions randomly within a specified percentage of their original values. The optimisers were then assigned the task of minimizing the difference in electrical responses between the perturbed structure and the original structure by varying the perturbed dimensions. This was typically repeated a number of times for each perturbation range. The results are plotted as a histogram of the resultant sum squared error obtained for each type of optimizer at each perturbation level. Figure A1.1 summarizes the results of an evaluation which is representative of the typical results obtained with this test.

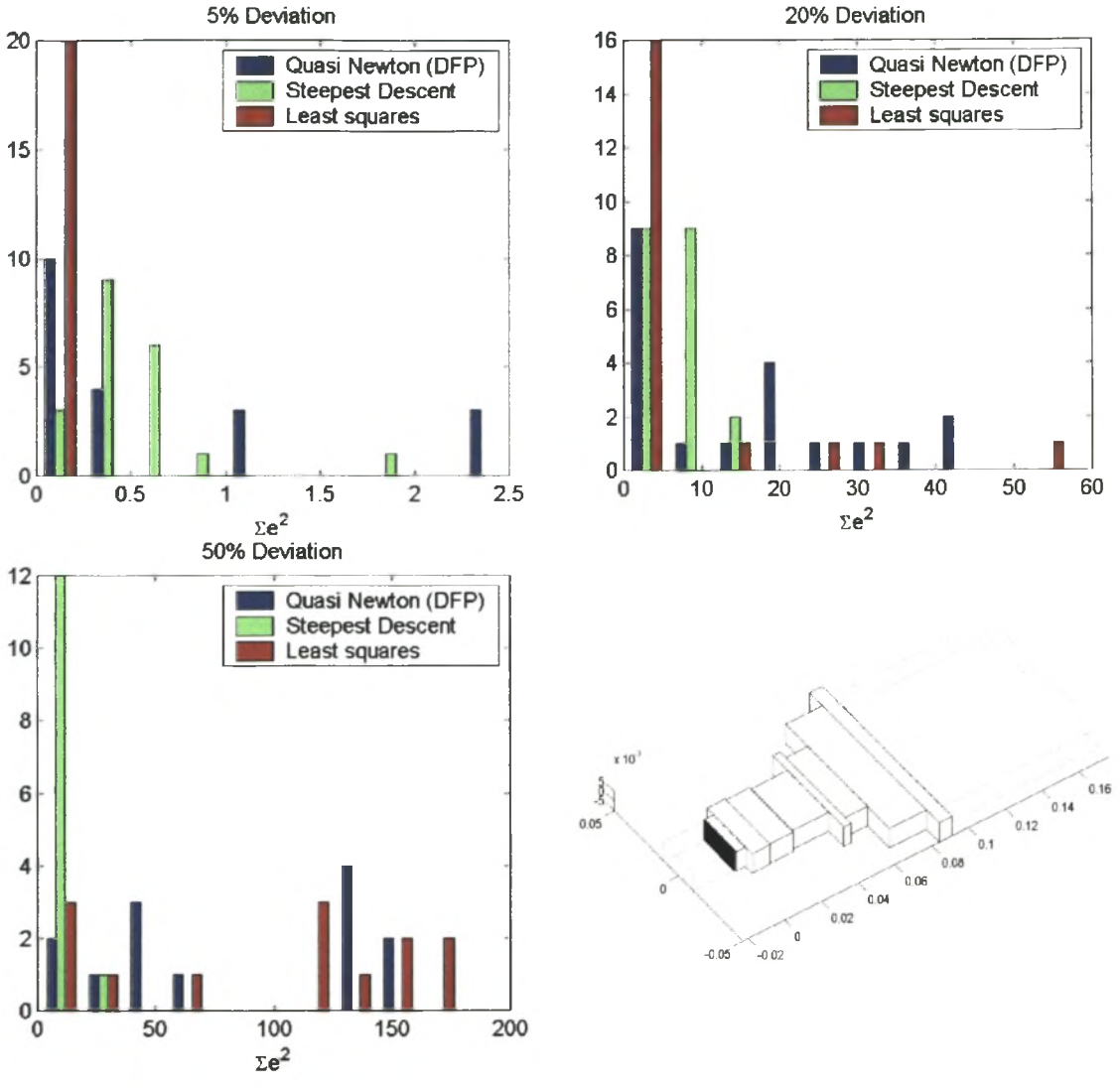


Figure A1.1 Relative performance of various gradient based optimisers for different perturbations of the test structure.

As well observed in the literature, the steepest descent technique provides excellent convergence far from the optimal point, but rather dismal performance near the minima. Separate tests with the Nelder-Mead Simplex technique, also shows excellent convergence at large deviations, but utilizes excessive function evaluations over the gradient based techniques.

Both the quasi-Newton and Least Squares methods were observed to struggle rather badly far away from the minima. Further investigation revealed that the error function (far from the minima) due to this type of problem does not conform well to the idealized paraboloid problem typically associated with mathematical optimization. Instead, it tends to produce

error functions that appear to have increasing gradients as the minima is approached, similar to the curve obtained by plotting the function, $y = \sqrt{|x|}$. This effect is generally less apparent nearer the local minima. It does, however, have the effect of causing the optimisers to 'over jump', resulting in unreliable convergence and wasted function evaluations.

Despite this, least squares is attractive as it is especially suited to error functions consisting of multiple outputs, which are widely utilized in this investigation. Steepest descent is attractive as it provides convergence approaching the reliability of simplex methods, at much less the cost in function evaluations. Accordingly, a hybridisation of least squares and steepest descent was implemented.

A1.3 Hybrid Steepest Descent – Least Squares Algorithm

The algorithm will not be detailed in full, as the least squares and steepest descent methods are well detailed in the literature [Polak 1997, Kelley 1999]. The following point form description of the algorithm details the essential components

- The optimizer calculates both the least squares and steepest descent search directions and adopts the direction which proves more favourable after testing each direction at a single point. For the least squares direction, this is the point predicted to be where the gradient is zero, and for the steepest descent direction, this is a point ¼ the distance of the previous iteration's jump.

- Gradients are calculated with finite differences.

- There are three termination criteria:
  - When the improvement in sum squared error over the last few iterations has been less than a specified minimum.
  - When the maximum number of iterations is exceeded.
  - When no further improvement can be made by jumping at least twice the value of finite difference step size used to calculate the gradient.

Before terminating, both the least squares and steepest descent direction must be fully investigated to ensure there is no viable jump to be found in either direction.

- In order to allowed constrained optimization, the error function may return a list of variables which have been truncated to adhere to maximum value of minimum value constraints. The optimizer then compares the returned variables to those it sent to the error function and modifies its finite difference step direction and current location accordingly, to account for the constraints.

The algorithm was tested with a number of simple examples in addition to the standard test discussed above and found to work reliably, at the cost a slightly more function evaluations than the individual techniques used alone. The cost of additional functional evaluations incurred by testing both steepest descent and lest squares directions at least once each iteration is minimal in comparison to the number if function evaluations typically involved in calculating the derivative information. This algorithm was used for most problems in this investigation requiring optimization, except in a few cases where standard least squares would suffice.

Appendix 2

Growth-based CAD Algorithms

BASIC SCANNING STRATEGY

0) *Definition of Constants*

NumRand	= Number of random starting seeds
MaxLength	= Maximum physical length of structure (set to infinity to leave unspecified)
FBs	= The set of available functional blocks
MinImprove	= The minimum improvement that will allow a FB inclusion in the structure

Define the input and output port dimensions

1) Initiate ***NumRand*** random starting structures with a minimum of functional blocks separated by appropriate random lengths of line.

2) Optimize each of the random starting structures.

3) Set ***CurrentStructure*** as the best outcome of the structure optimized in setp 2) and set E^{Curr} as its sum squared error

4) Set E^{Best} to infinity.

5) Loop through all viable insertion points, ***p=1..N***, in ***CurrentStructure***

5.1) Loop through all available functional block types, ***FB(q)***

5.1.1) Initiate an new random instance of ***FB(q)*** and place it in ***CurrentStructure*** at location ***p*** by splitting the existing line at ***p*** in two and adding random lengths to the lines. If ***p*** is at the beginning or end of the structure, add a random length line between the beginning or end of the structure and the new functional block.

5.1.2) Optimize the test structure by varying only the design variables corresponding to the newly inserted functional block and the adjacent line lengths.

5.1.3) Now optimize the entire test structure by varying all the available design variables and obtain the resultant structure's sum squared error, E^{Test}.

5.1.4) Obtain the physical length of the test structure, ***TestLength***, from the Rubberband algorithm.

5.1.5) If ***TestLength*** < ***MaxLength*** and $E^{Test} < E^{Best}$,

- Set $E^{Best} = E^{Test}$,
- Store the test structure as ***NewStructure***

6) If $(E^{Curr} - E^{Best}) >$ ***MinImprove***,

- Set $E^{Curr} = E^{Best}$
- Set ***CurrentStructure*** = ***NewStructure***

Else

- Unsuccessful scan

7) *Reserved for Redundant Line Remover*

8) *Reserved for Redundant Functional Block Remover*

9) *Termination criteria*

- Number of functional blocks in ***CurrentStructure*** is greater than a specific maximum
- Number of unsuccessful scans in a row is greater than a specific maximum
- E^{Curr} is less than a specified minimum

else

Go to step 4) for a new scan

REDUNDANT LINE REMOVER

1) Loop through all connecting line sections, ***p = 1..N***

1.1) Obtain the modes that are excited in connecting section *p* of ***CurrentStructure*** from the analysis technique.

1.2) If only one of these modes is propagating in section ***p***,

1.2.1) Determine the wavelength, λ, of the single propagating mode.

1.2.2) If the length of connecting section, ***p***, is larger than ½λ, successively remove ½λ lengths until the length of the connecting section is less than ½λ.

1.2.3) Optimize the entire modified structure and obtain the sum squared error, E^{Test} of the resultant structure, ***NewStructure***.

1.2.4) if ($E^{Test} < E^{Curr}$) OR ($E^{Test} - E^{Curr}$) < ***½ MinImprove,***

- Set $E^{Curr} = E^{Test}$
- Set ***CurrentStructure*** = ***NewStructure***

REDUNDANT FB REMOVER

1) Set E^{Best} to infinity.

2) Loop through all functional blocks, ***p = 1..N-1***, in ***CurrentStructure***.

2.1) Assign ***TestStructure*** = ***CurrentStructure***

2.2) Remove functional block, ***p***, from ***TestStructure*** and replace the adjacent line sections on either side of functional block p with a new random length line section.

2.3) Optimize ***TestStructure*** and obtain its resultant sum squared error, E^{Test}.

2.4) Obtain the physical length of the test structure, ***TestLength***, from the Rubberband algorithm.

2.5) ***TestLength < MaxLength*** and $E^{Test} < E^{Best}$,

- Set $E^{Best} = E^{Test}$,
- Store the test structure as ***NewStructure***

3) If ($E^{Best} < E^{Curr}$) OR ($E^{Best} - E^{Curr}$) < ***½ MinImprove,***

- Set $E^{Curr} = E^{Best}$
- Set ***CurrentStructure*** = ***NewStructure***

SIMPLIFIED TREE SEARCH STRATEGY (RECURSIVE)

0) *Definition of Constants*

NumRand	= Number of random starting seeds
MaxLength	= Maximum physical length of structure (set to infinity to leave unspecified)
FBs	= The set of available functional blocks
MinImprove	= The minimum improvement that will allow a FB inclusion in the structure

Define the input and output port dimensions

1) If this is the first instance of this algorithm,

True 1.1) Initiate ***NumRand*** random starting structures with a minimum of functional blocks separated by appropriate random lengths of line. The structure may consist of just one functional blocks and both ports, in which case ***NumRand*** is typically set to the number of Functional Blocks in ***FBs***

True 1.2) If appropriate, optimize the random structures and place the resultant structures in a set ***1..NumRand*** called ***CurrentStructures***.

OTHERWISE

False 1.1) Retrieve the Structure passed to this instance by the previous instance, ***OriginalStructure*** and its sum squared error, E^{Orig}.

False 1.2) Loop through all Functional Blocks ***q=1..M***, in ***FBs***,

False 1.2.1) Add the Functional Block, ***FBs(q)***, to the end of ***OriginalStructure***, separated by a length of random line, or line length determined from prior knowledge of the problem.

False 1.2.2) If appropriate, optimize the new structure and place the result in a set of ***CurrentStructures***.

2) Set a set of output structures, ***FinishedStructures*** to empty

3) Loop through all **CurrentStructures**, ***q=1..M***

3.1 Obtain the sum squared error, $E^{Cur}{}_{q}$ of ***CurrentStructures(q)***

3.2 Obtain the physical length, ***qLength***, of ***CurrentStructure(q)*** from the Rubberband Algorithm.

3.3 If ***(EOrig – ECurrq) > MinImprove*** AND ***qLength < MaxLength***,

- Call a new instance of this algorithm. Pass to it ***CurrentStructures(q)***, and add the returned output of this call to the ***FinishedStructures*** set

4) If FinshedStructures is empty (none of ***CurrentStructures*** was better than ***OriginalStructure***),

- Place ***OriginalStructure*** into ***FinsihedStructures***

5) Return ***FinishedStructures*** to the original calling function as output

EVOLTIONARY STRATEGY

0) *Definition of Constants*

NumPop	= Population Size
NumGen	= Number of Generations
MaxLength	= Maximum physical length of structure (set to infinity to leave unspecified)
FBs	= The set of available functional blocks

Define the input and output port dimensions

1) Initiate ***NumPop*** random starting structures, ***Pop(1..NumPop)***, with a minimum of functional blocks separated by appropriate random lengths of line. The structure may consist of just one functional blocks and both ports.

2) Analyze the entire population, ***Pop,*** and acquire their individual sum squared errors, $E^{Pop}(1..NumPop)$.

3) Loop through the number of generations, ***p=1..NumGen***,

3.1) Determine best 50% of the ***Pop*** (SEE "Competition" BELOW)

3.2) Produce ***½NumPop*** new offspring from the best 50% of ***Pop*** (SEE "Reproduction" BELOW)

3.3) Analyze the new 50% of ***Pop***, and acquire their individual sum squared errors.

EVOLUTIONARY STRATEGY – COMPETITION

1) Loop from 1 to ***½NumPop,***

1.1) Take 2 random population members, ***Pop(p)*** and ***Pop(q)***

1.2) Find the physical lengths of ***Pop(p)***, ***pLength***, and ***Pop(q)***, ***qlength***.

1.3) if ***EPop(p) >> EPop(q)***,

- Select ***EPop(q)*** as the winner

1.4) if ***EPop(q) >> EPop(p)***,

- Select ***EPop(p)*** as the winner

1.5) if ***EPop(q) ≈ EPop(p)***,

- Select the winner corresponding to the lowest of ***qLength*** and ***pLength***

1.6) Save the winner into the new population for the next generation.

EVOLUTIONARY STRATEGY – REPRODUCTION

1) Loop from 1 to ***½NumPop,***

1.1) Take 2 random population members from the new population, ***Pop(p)*** and ***Pop(q)***

1.2) Decide to pursue sexual or asexual reproduction base on the average size of the parents. If the size is small, favour asexual reproduction, otherwise favour sexual reproduction (randomly)

1.2.1) Pursue sexual reproduction by splitting the two parent structures at a joining line section at random point on their length and joining two random halves from both parents with a new line section.

OR

1.2.2) Pursue asexual reproduction by randomly selecting one of the parents and either, **a)** Adding a functional block, **b)** removing a functional block (if possible), **c)** moving a functional block, **d)** modifying a functional block or line section's design variables.

1.3) If the offspring is not viable (i.e. not physically possible) or has a length longer than ***MaxLength***, discard the offspring and retry with new parents until a viable structure is found.

Appendix 3

Normalization of Modal Power

The modal fields must be normalized to carry unity magnitude power. This facilitates the interpretation of the results of mode matching in terms of the *Generalized Scattering Matrix* formulation. The modal fields as derived in [Ramo 1993b] are repeated in eqns. (A3.1) and (A3.2).

TM Fields	**TE Fields**
$E_x = C_{TM}\left[-\frac{\gamma k_x}{k_c^2}\cos(k_x x)\sin(k_y y)\right]$	$E_x = C_{TE}\left[\frac{j\omega\mu k_y}{k_c^2}\cos(k_x x)\sin(k_y y)\right]$
$E_y = C_{TM}\left[-\frac{\gamma k_y}{k_c^2}\sin(k_x x)\cos(k_y y)\right]$	$E_y = C_{TE}\left[-\frac{j\omega\mu k_x}{k_c^2}\sin(k_x x)\cos(k_y y)\right]$
$E_z = C_{TM}\left[\sin(k_x x)\cos(k_y y)\right]$	$E_z = 0$
$H_x = C_{TM}\left[\frac{j\omega\varepsilon k_y}{k_c^2}\sin(k_x x)\cos(k_y y)\right]$	$H_x = C_{TE}\left[\frac{\gamma k_x}{k_c^2}\sin(k_x x)\cos(k_y y)\right]$
$H_y = C_{TM}\left[-\frac{j\omega\varepsilon k_x}{k_c^2}\cos(k_x x)\sin(k_y y)\right]$	$H_y = C_{TE}\left[\frac{\gamma k_y}{k_c^2}\cos(k_x x)\sin(k_y y)\right]$
$H_z = 0$	$H_z = C_{TE}\left[\cos(k_x x)\sin(k_y y)\right]$

(A3.1)

Wave Number and Propagation Constant

$$\gamma^2 = k_c^2 - k^2 \qquad k_x = \frac{m\pi}{a}$$

$$k_c^2 = k_x^2 + k_y^2 \qquad k_y = \frac{n\pi}{b} \tag{A3.2}$$

Eqn (A3.3) is now investigated for the TE and TM modes in turn.

$$\mathrm{P} = \iint_{S_{WG}} \left[\left(\mathbf{E}\times\mathbf{H}^*\right)\cdot\hat{i}_z\right]\cdot dA \tag{A3.3}$$

A3.1 TE Modes

$$P = \int_0^b \int_0^a |C_{TE}|^2 \left[\frac{j\omega\mu k_y \gamma^* k_y}{k_c^4} \cos^2(k_x x) \sin^2(k_y y) + \frac{j\omega\mu k_x \gamma^* k_x}{k_c^4} \sin^2(k_x x) \cos^2(k_y y) \right] \cdot dxdy$$

$$\therefore P = |C_{TE}|^2 \frac{j\omega\mu k_c^2 \gamma^* ab}{k_c^4} \frac{1}{q^2} \qquad \text{with } q = \begin{cases} \sqrt{2} & \text{for } m = 0 \text{ or } n = 0 \\ 2 & \text{Otherwise} \end{cases} \tag{A3.4}$$

Now, with $|P| = 1$, this specifies,

$$\therefore |C_{TE}| = \sqrt{\frac{k_c^2 q^2}{\omega\mu|\gamma|ab}}$$

From eqns. (A3.4), it is clear that the choice of phase of C_{TE} is arbitrary. One option is to choose C_{TE} as purely real as in eqn. (A3.5). An alternative choice, shown in eqn. (A3.6) is equally valid and when substituted in eqns. (A3.1) for the TE modes results in eqns. (A3.7).

$$C_{TE} = \frac{k_c q}{\sqrt{ab}} \frac{1}{\sqrt{|\gamma|}} \frac{1}{\sqrt{\omega\mu}} \tag{A3.5}$$

$$C_{TE} = \frac{k_c q}{\sqrt{ab}} \frac{1}{\sqrt{j\gamma\omega\mu}} \tag{A3.6}$$

$$\begin{aligned}
E_x &= \frac{k_c q}{\sqrt{ab}} \frac{1}{\sqrt{j\gamma\omega\mu}} \frac{j\omega\mu k_y}{k_c^2} C_x S_y = \frac{q k_y}{k_c \sqrt{ab}} \sqrt{\frac{j\omega\mu}{\gamma}} C_x S_y = \frac{q}{\sqrt{ab}} \sqrt{Z_{TE}} \frac{k_y}{k_c} C_x S_y \\
E_y &= \frac{k_c q}{\sqrt{ab}} \frac{1}{\sqrt{j\gamma\omega\mu}} \frac{j\omega\mu k_x}{k_c^2} S_x C_y = \frac{q k_x}{k_c \sqrt{ab}} \sqrt{\frac{j\omega\mu}{\gamma}} S_x C_y = \frac{q}{\sqrt{ab}} \sqrt{Z_{TE}} \frac{k_x}{k_c} S_x C_y \\
H_x &= \frac{k_c q}{\sqrt{ab}} \frac{1}{\sqrt{j\gamma\omega\mu}} \frac{\gamma k_x}{k_c^2} S_x C_y = \frac{q k_x}{k_c \sqrt{ab}} \sqrt{\frac{\gamma}{j\omega\mu}} S_x C_y = \frac{q}{\sqrt{ab}} \sqrt{Y_{TE}} \frac{k_x}{k_c} S_x C_y \\
H_y &= \frac{k_c q}{\sqrt{ab}} \frac{1}{\sqrt{j\gamma\omega\mu}} \frac{\gamma k_y}{k_c^2} C_x S_y = \frac{q k_y}{k_c \sqrt{ab}} \sqrt{\frac{\gamma}{j\omega\mu}} C_x S_y = \frac{q}{\sqrt{ab}} \sqrt{Y_{TE}} \frac{k_y}{k_c} C_x S_y
\end{aligned} \tag{A3.7}$$

The normalized quantities in eqns. (A3.7) can be checked by substituting back in eqn. (A3.3). Note that since the quantities in eqns. (A3.7) can incorporate square roots of complex quantities, care must be taken to use the same choice of solution to the square root consistently. This is typically accomplished by the numerical algorithm used for the square root computation. MATLAB always uses the solution with positive real component and this choice suffices for this formulation.

A3.2 TM Modes

$$P = \int_0^b \int_0^a |C_{TM}|^2 \left[-\frac{\gamma k_x j\omega\varepsilon k_x}{k_c^4} \cos^2(k_x x)\sin^2(k_y y) - \frac{\gamma k_y j\omega\varepsilon k_y}{k_c^4} \sin^2(k_x x)\cos^2(k_y y) \right] \cdot dxdy$$

$$\therefore P = |C_{TM}|^2 \frac{-j\omega\varepsilon k_c^2 \gamma ab}{k_c^4} \frac{1}{4} \tag{A3.8}$$

Now, with $|P| = 1$, this specifies,

$$\therefore |C_{TM}| = \sqrt{\frac{4k_c^2}{\omega\varepsilon|\gamma|ab}}$$

A similar option to that chosen for the TE modes, shown in eqn. (A3.9), is selected and substituted in eqns (A3.1) for the TM modes.

$$C_{TM} = \frac{2k_c}{\sqrt{ab}} \frac{1}{\sqrt{j\gamma\omega\varepsilon}} \tag{A3.10}$$

$$\begin{aligned}
E_x &= -\frac{2k_c}{\sqrt{ab}} \frac{1}{\sqrt{j\gamma\omega\varepsilon}} \frac{\gamma k_x}{k_c^2} C_x S_y = -\frac{2k_x}{k_c\sqrt{ab}} \sqrt{\frac{\gamma}{j\omega\varepsilon}} C_x S_y = -\frac{2}{\sqrt{ab}} \sqrt{Z_{TM}} \frac{k_x}{k_c} C_x S_y \\
E_y &= -\frac{2k_c}{\sqrt{ab}} \frac{1}{\sqrt{j\gamma\omega\varepsilon}} \frac{\gamma k_y}{k_c^2} S_x C_y = \frac{2k_y}{k_c\sqrt{ab}} \sqrt{\frac{\gamma}{j\omega\varepsilon}} S_x C_y = \frac{2}{\sqrt{ab}} \sqrt{Z_{TM}} \frac{k_y}{k_c} S_x C_y \\
H_x &= \frac{2k_c}{\sqrt{ab}} \frac{1}{\sqrt{j\gamma\omega\varepsilon}} \frac{j\omega\varepsilon k_y}{k_c^2} S_x C_y = \frac{2k_y}{k_c\sqrt{ab}} \sqrt{\frac{j\omega\varepsilon}{\gamma}} S_x C_y = \frac{2}{\sqrt{ab}} \sqrt{Y_{TM}} \frac{k_y}{k_c} S_x C_y \\
H_y &= -\frac{2k_c}{\sqrt{ab}} \frac{1}{\sqrt{j\gamma\omega\varepsilon}} \frac{j\omega\varepsilon k_x}{k_c^2} C_x S_y = -\frac{2k_x}{k_c\sqrt{ab}} \sqrt{\frac{j\omega\varepsilon}{\gamma}} C_x S_y = -\frac{2}{\sqrt{ab}} \sqrt{Y_{TM}} \frac{k_x}{k_c} C_x S_y
\end{aligned} \tag{A3.11}$$

As before, the normalized quantities in eqns. (A3.11) can be verified by substituting back in eqn. (A3.3).

A3.3 Non-unique GSM formulations

Both the choice of phase of C_{TE} and C_{TM} and the choice of square root solution indicate that, in general, the GSM formulation is not unique. In practice this results in some off-diagonals of the GSM matrices having non-unique phases, depending on the choice of normalization. The choice made above corresponds to that made by Alessandri et al. [Alessandri 1992].

When cascading GSMs it is however critical that they have the same normalization convention. This is inherently satisfied in this investigation by the fact that only the mode matching formulation derived in chapter 3 shall be used to solve for GSMs. In hybrid techniques, which might cascade GSMs from many different sources (e.g. MoM, MM, FEM), care must be taken to ensure that the choice of normalization used for all the different techniques is the same.

To confirm that the choice of normalization in the MM formulation does not lead to conflicting scattering parameters of cascaded structures (as long as the choice is used consistently), an alternative MM formulation, which used zero-phase C_{TE}'s and C_{TM}'s as suggested in eqn. (A3.5) was also developed. A number of sample cascaded step structures were analysed with both the standard formulation and the alternative formulation. As expected, the final GSMs of agreed in all but the phase of some of the off-diagonals.

These discrepancies must be taken into consideration when total EM fields are to be determined. In Chapter 6, where the modal weighting is to be determined to shape field patterns, the procedure that determines the total field quantities must be formulated in terms of the correct normalization convention. In section 3.7, where transforms between modal sets are derived, the issue of modal power normalization arises again. In that case, the derivation was done with the more standard zero-phase C_{TE} and C_{TM} choice. The transforms that result, however, appear to be normalization convention independent.

Appendix 4

Solution to the Problem of ill-conditioning in T-Parameter Analysis of Cascaded Structures

A4.1 Introduction

The use of the transmission parameters (T-parameters) for analysis of cascaded waveguide structures has been limited due to its undesirable numerical properties when dealing with non-propagating modes. Numerous texts cite this as the reason for choosing the scattering parameter formulation when solving cascaded systems, despite the increased numerical effort involved (See various textbooks on T-parameters). The T-parameters, however, offer significant advantages. Firstly, cascading requires only a matrix multiplication, as opposed to an obligatory matrix inversion required by other formulations. Secondly, the T-parameter representation of a discontinuity solved with mode matching requires only one *frequency independent* real matrix inversion with minimal additional matrix operations. This is opposed to the complex matrix inversion *and* considerable additional matrix operations required *per frequency point* for the S-parameter formulation. This implies that for complete cascade system analysis (with mode matching) over an arbitrary number of frequency points, T-parameters require only N real matrix inversions, where N is the total number of discontinuities. Another advantage of the T-matrix formulation is the similarity in its use for synthesis and analysis.

One approach that attempts to exploit the advantages of the T-parameters is to apply the S-parameters to substructures which have been cascaded with the T-parameters. The drawback of this approach is that its efficiency tends to drop with complex structures, or when significant accuracy is required (many modes).

This proposes a new approach that stabilizes the T-parameter formulation, allowing full T-parameter cascading with all its inherent advantages. The method restricts the attenuation of modes and consequently the magnitudes of the T-parameters by incorporating 'helping' modes to provide additional signal paths for the highly attenuated modes.

This text will start by briefly discussing the reasons for T-parameter instability in cascade waveguide discontinuity analysis. It will then introduce the stabilizing operator and finish with comparisons between calculated parameters of a sample device using the new method and the traditional scattering parameter method.

A4.2. T-Parameter Instability

The T-parameter formulation, expressed in terms of wave amplitudes, can be summarized as in (A4.1)

$$\begin{bmatrix} \bar{a}_A \\ \bar{b}_A \end{bmatrix} = \begin{bmatrix} [A] & [B] \\ [C] & [D] \end{bmatrix} \begin{bmatrix} \bar{b}_B \\ \bar{a}_B \end{bmatrix} = [T] \begin{bmatrix} \bar{b}_B \\ \bar{a}_B \end{bmatrix} \tag{A4.1}$$

Here $[T]$ is the *wave* transmission matrix and $\bar{a}$ and $\bar{b}$ are complex amplitudes of the incident and reflected waves respectively on physical ports A and B of the device. The T-parameters and S-parameters relate to each other as in (A4.2) and (A4.3)

$$\begin{aligned} [S_{AA}] &= [C][A]^{-1} \\ [S_{AB}] &= -[C][A]^{-1}[B]+[D] \\ [S_{BA}] &= [A]^{-1} \\ [S_{BB}] &= -[A]^{-1}[B] \end{aligned} \tag{A4.2}$$

$$\begin{aligned} [A] &= [S_{BA}]^{-1} \\ [B] &= -[S_{BA}]^{-1}[S_{BB}] \\ [C] &= [S_{AA}][S_{BA}]^{-1} \\ [D] &= [S_{AB}]-[S_{AA}][S_{BA}]^{-1}[S_{BB}] \end{aligned} \tag{A4.3}$$

Consider now a system of 2 waveguide modes (4 electrical ports) in which the second mode is greatly attenuated in transmission, i.e. $S_{BA}(2,2) \approx 0$, and some cross-coupling occurs between the modes. In this case, the resultant $[A]$ matrix will have a very large value at $A(2,2)$. When such systems are

cascaded, the extremely large numerical values will cause loss of numerical accuracy, due to truncation error in the machine, and consequently loss of linear independence of the $[A]$ matrix columns.

This problem often arises in cascaded waveguide discontinuity problems when non-propagating modes must traverse appreciable lengths of connecting waveguide separating the discontinuities. A full *scattering* parameter analysis of such a system would result in a $[S_{BA}]$ matrix whose diagonal elements corresponding to the strongly non-propagating modes would be very close to zero, implying a potentially badly conditioned $[A]$ matrix. The reason for this ill-conditioned behaviour can been seen from the generalized T-parameters for a length of transmission line, l, with traveling modes of propagation constant, $\gamma = \alpha + j\beta$, as given in eqn. (A4.4). Clearly, the exponent raised to a positive real quantity in the $[A^{line}]$ sub-matrix can result in very large values for strongly non-propagating modes.

$$[T^{line}] = \left[\begin{array}{ccc|ccc} e^{\gamma_1 l} & 0 & 0 & & & \\ 0 & e^{\gamma_2 l} & 0 & & [0] & \\ 0 & 0 & \ldots & & & \\ \hline & & & e^{-\gamma_1 l} & 0 & 0 \\ & [0] & & 0 & e^{-\gamma_2 l} & 0 \\ & & & 0 & 0 & \ldots \end{array}\right] \tag{A4.4}$$

A4.3. Stabilized T-Parameters

In order to address the problem of numerical instability in T-parameter cascading, it is necessary to place an upper bound on the magnitudes of the T-parameters caused by strong attenuation in the structure. To this end, the concept of 'helping' modes is proposed here. Helping modes are non-physical, non-interacting modes with no phase change but with a specified attenuation, temporarily incorporated in such a structure to provide alternate signal paths for the unmanageable modes.

As discontinuities and line lengths are cascaded, the T-parameters will show increasing instability for the non-propagating modes. As this occurs, the magnitudes of the $[A]$ matrix entries corresponding to

these modes increase, indicating their increasing attenuation. When a particular mode's attenuation exceeds a specific limit imposed by the machine's numerical range, r, the cascaded parameters must be stabilized. This is accomplished by placing scatterers on either side of the partly cascaded system that scatter energy to a helping mode. The helping mode is attenuated by v, where v is some orders of magnitude less than r.

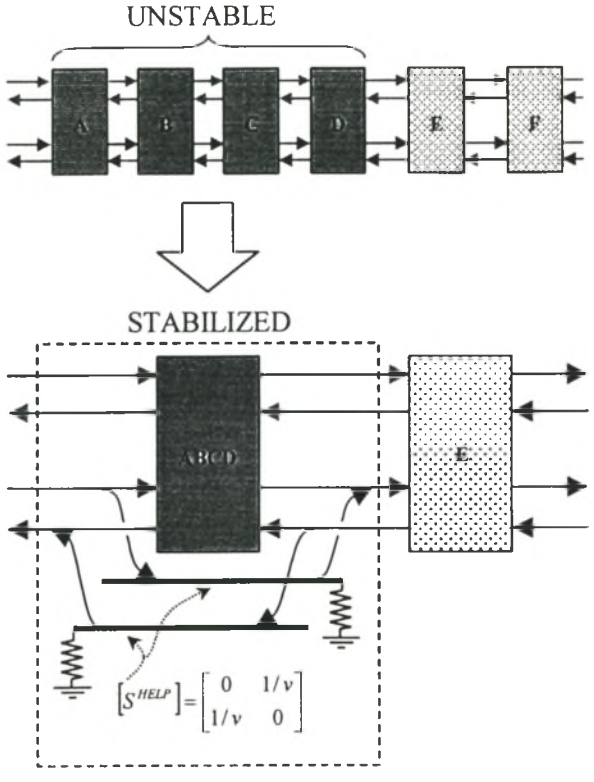


Figure A4.1. The stabilizing operation

In figure A4.1, the procedure has been applied to the partly cascaded system ABCD, having done this, it might then have to be reapplied for the partly cascaded system, (ABCD)EFG and so on until the cascading is completed. The attenuation of the problematic modes is in this way effectively maintained at a level low enough to ensure numerical stability. Note that the choice of the helping mode attenuation, v, must be set sufficiently high to maintain accuracy of the cascading operation. If v is chosen closer to r, the limit of stability imposed by the machine, the error introduced by the helping mode to the complete system parameters will decrease but the stabilizing process will have to occur more frequently to ensure that the maximum attenuation stays below r.

A4.4 Implementing Stabilization

Figure A4.2 shows a block diagram of the stabilization operation for the two mode example.

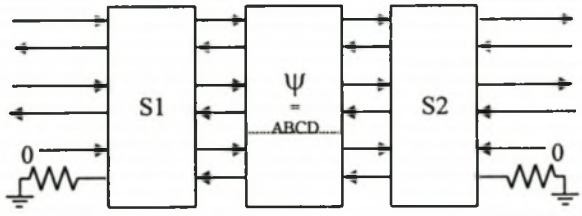


Figure A4.2. Block diagram of stabilization

The scattering parameters of the extended system, Ψ (including the helping mode) together with its T-parameter equivalent are given in eqns. (A4.5) and (A4.6) in terms of the S-parameters of the system to be stabilized.

$$[S^{\Psi}]=\left[\begin{array}{cc:cc} [S^{\Psi}_{AA}] & \bar{0}^T & [S^{\Psi}_{AB}] & \bar{0}^T \\ \bar{0} & 0 & \bar{0} & 0 \\ \hdashline [S^{\Psi}_{BA}] & \bar{0}^T & [S^{\Psi}_{BB}] & \bar{0}^T \\ \bar{0} & 0 & \bar{0} & 0 \end{array}\right] \qquad \text{(A4.5)}$$

$$[T^{\Psi}]=\left[\begin{array}{cc:cc} [A^{\Psi}] & \bar{0}^T & [B^{\Psi}] & \bar{0}^T \\ \bar{0} & v & \bar{0} & 0 \\ \hdashline [C^{\Psi}] & \bar{0}^T & [D^{\Psi}] & \bar{0}^T \\ \bar{0} & 0 & \bar{0} & 1/v \end{array}\right] \qquad \text{(A4.6)}$$

Here $\bar{0}$, is the nil vector, [0 0 …].

The scattering parameters of the scatterers, S1 and S2 follow in eqns. (A4.7) and (A4.9) with their respective T-parameter equivalents in (A4.8) and (A4.10).

$$[S^{S1}]=\left[\begin{array}{c:c} [0] & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \\ \hdashline \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} & [0] \end{array}\right] \qquad \text{(A4.7)}$$

$$[T^{S1}]=\left[\begin{array}{c:c} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} & [0] \\ \hdashline [0] & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \end{array}\right] \qquad \text{(A4.8)}$$

$$[S^{S2}]=\left[\begin{array}{c|c} [0] & \begin{bmatrix}1&0&0\\0&1&0\\0&1&1\end{bmatrix} \\ \hline \begin{bmatrix}1&0&0\\0&1&1\\0&0&1\end{bmatrix} & [0] \end{array}\right] \qquad \text{(A4.9)}$$

$$[T^{S2}]=\left[\begin{array}{c|c} \begin{bmatrix}1&0&0\\0&1&-1\\0&0&1\end{bmatrix} & [0] \\ \hline [0] & \begin{bmatrix}1&0&0\\0&1&0\\0&1&1\end{bmatrix} \end{array}\right] \qquad \text{(A4.10)}$$

Having cascaded in the scatterers, S1 and S2, the helping mode must be terminated and the complete system rendered a 4 port once again. Terminating with a matched load is trivial in S-parameters – the T-parameter equivalent operation is shown in eqns. (A4.11) and (A4.12) for M modes. Consider a T-parameter system as shown in (A4.11), with M ports per side where $M = N+R$. R are the ports to be terminated.

$$\left[\begin{array}{c} \bar{a}_A^N \\ \bar{a}_A^R \\ \hline \bar{b}_A^N \\ \bar{b}_A^R \end{array}\right]=\left[\begin{array}{cc|cc} [A^{NN}] & [A^{NR}] & [B^{NN}] & [B^{NR}] \\ [A^{RN}] & [A^{RR}] & [B^{RN}] & [B^{RR}] \\ \hline [C^{NN}] & [C^{NR}] & [D^{NN}] & [D^{NR}] \\ [C^{RN}] & [C^{RR}] & [D^{RN}] & [D^{RR}] \end{array}\right] \qquad \text{(A4.11)}$$

The new 2N port network T-parameters can be expressed as in (A4.12).

$$\begin{aligned} [A^{eq}] &= [A^{NN}]-[A^{NR}][A^{RR}]^{-1}[A^{RN}] \\ [B^{eq}] &= [B^{NN}]-[A^{NR}][A^{RR}]^{-1}[B^{RN}] \\ [C^{eq}] &= [C^{NN}]-[C^{NR}][A^{RR}]^{-1}[A^{RN}] \\ [D^{eq}] &= [D^{NN}]-[C^{NR}][A^{RR}]^{-1}[B^{RN}] \end{aligned} \qquad \text{(A4.12)}$$

Note that when terminating only a single mode, this operation requires no inversion.

Due to the simple nature of the scatterers and the helping mode, the entire stabilizing operation can be very easily accomplished using only a few simple matrix element operations. The procedure applied to general multimode cascading follows simply from the two-mode example above.

A4.5 Results

The stabilized T-parameter analysis was applied to solving a sample cascaded structure in waveguide that cannot be solved for using conventional T-parameters. Figure. A4.3 details some results of this analysis together with results of analysis done with the standard S-parameter technique. For the purposes of comparison, the resultant T-parameters have been converted to S-parameters. Note the limit in dynamic range imposed by the v parameter in the S_{21} of the TE_{20} mode.

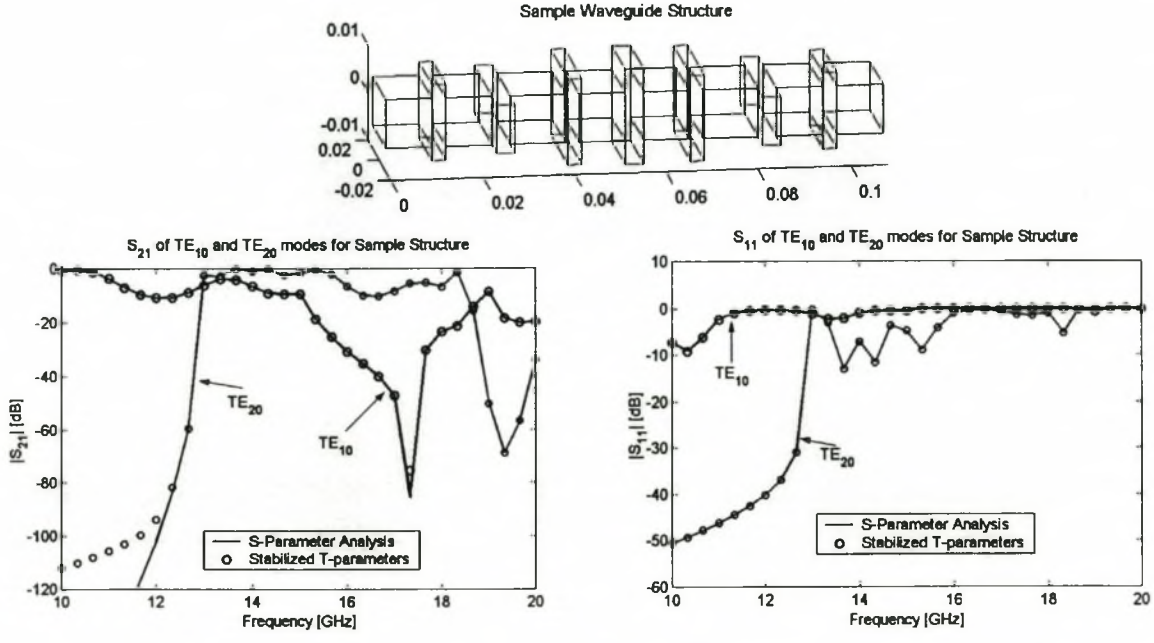


Figure. A4.3 Sample structure analysis with about 100 modes. Dimensions are in metres

A4.6 Conclusion

A technique of stabilizing the T-parameters for the purpose of efficiently cascading waveguide discontinuities has been presented. The technique relies on the use of helping modes to maintain the attenuation of badly conditioned ports to below a critical value. Results are presented that show good agreement with standard S-parameter analysis techniques.