

**A UNIFIED, INTEGRATED APPROACH  
TO GENERALIZED  $\ell_p$  CIRCUIT OPTIMIZATION**

S.H. Chen

SOS-87-3-T

September 1987

© S.H. Chen 1987

No part of this document may be copied, translated, transcribed or entered in any form into any machine without written permission. Address enquiries in this regard to Dr. J.W. Bandler. Excerpts may be quoted for scholarly purposes with full acknowledgement of source. This document may not be lent or circulated without this title page and its original cover.



**A UNIFIED, INTEGRATED APPROACH TO  
GENERALIZED  $\ell_p$  CIRCUIT OPTIMIZATION**



*To my wife Bei Lu*



**A UNIFIED, INTEGRATED APPROACH TO  
GENERALIZED  $\ell_p$  CIRCUIT OPTIMIZATION**

By

SHAOHUA CHEN, B.S.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

September 1987



DOCTOR OF PHILOSOPHY (1987)  
(Electrical Engineering)

McMASTER UNIVERSITY  
Hamilton, Ontario

**TITLE:** A Unified, Integrated Approach to Generalized  $\ell_p$   
Circuit Optimization

**AUTHOR:** Shaohua Chen, B.S. (Eng.)  
(South China Institute of Technology)

**SUPERVISOR:** J.W. Bandler, Professor, Department of Electrical and  
Computer Engineering  
B.Sc.(Eng.), Ph.D., D.Sc.(Eng.) (University of London)  
D.I.C., (Imperial College)  
P.Eng. (Province of Ontario)  
C.Eng., F.I.E.E. (United Kingdom)  
Fellow, I.E.E.E.  
Fellow, Royal Society of Canada

**NUMBER OF PAGES:** xv, 221



## ABSTRACT

This thesis offers a unified and integrated treatment of three essential aspects of computer-aided circuit design: effective use of the state-of-the-art optimization tools, efficient calculation of exact and approximate gradients, and adequate mathematical representation of the engineering problems.

The recent advances in gradient-based  $\ell_p$  optimization are reviewed. The essence of the trust region Gauss-Newton method and the quasi-Newton solution to optimality equations is described. A new algorithm for linearly constrained one-sided  $\ell_1$  optimization is presented.

Efficient approaches to network sensitivity analysis are addressed. Useful formulas are derived for general multi-ports, especially two-ports. Novel proofs of an important result for lossless two-ports are given.

The basic formulations of nominal circuit optimization are introduced through a hierarchy of simulation models. Variables, error functions and  $\ell_p$  objectives are identified. Optimization of multi-coupled cavity filters is described and illustrated by examples of elliptic, self-equalized and asymmetric designs. Large-scale optimization of multiplexers is also discussed.

Realistic consideration of tolerances and uncertainties is of prominent interest to circuit, especially integrated circuit, designers. A multi-circuit approach to design centering, toleran-

cing, tuning and yield enhancement is presented. Techniques for statistical design are reviewed. A generalized  $\ell_p$  centering algorithm is developed.

A novel approach to device modeling which utilizes multiple circuits and exploits the theoretical properties of the  $\ell_1$  norm is described. It emphasizes the uniqueness and consistency of an equivalent circuit model. Practical applications are formulated and illustrated through industrial examples.

A new algorithm for optimization with integrated gradient approximations is offered. Implementations for the minimax and  $\ell_1$  problems are shown. The efficiency and usefulness are demonstrated by a large variety of examples.

### ACKNOWLEDGEMENTS

The author wishes to express his sincere appreciation to Dr. J.W. Bandler for his encouragement, expert guidance and supervision throughout the course of this work. He also thanks Dr. C.M. Crowe and Dr. N.K. Sinha, members of his Supervisory Committee, for their continuing interest.

The author appreciates the opportunity given to him by Dr. Bandler to be involved in industrially relevant projects. Cooperation between Dr. Bandler and Dr. R.A. Pucel of Raytheon Company, Lexington, MA, was a significant motivation for the work on statistical design. Cooperation with ComDev Ltd. of Cambridge, Ontario, in preparing multi-coupled cavity filter data is acknowledged. Some FET measurement data was kindly made available by Raytheon Company.

The author offers his deep gratitude to Dr. K. Madsen of the Technical University of Denmark for his expert advice, through lectures, private communications and computer programs, on optimization theory and gradient approximations.

It is the author's pleasure to have been associated with Dr. S. Daijavad. The continuous and in-depth interaction with Dr. Daijavad has led to the development of new ideas and applications. The author has also benefitted from inspiring discussions with his colleagues M.L. Renault, Q.J. Zhang and Dr. W. Kellermann.

The financial assistance provided by the Ministry of Education of the People's Republic of China through an overseas graduate scholarship, the Natural Sciences and Engineering Research Council of Canada through Grants A7239 and G1135, the Department of Electrical and Computer Engineering through a Teaching Assistantship and the Ministry of Colleges and Universities through an Ontario Graduate Scholarship is gratefully acknowledged.

## TABLE OF CONTENTS

	PAGE
<b>ABSTRACT</b>	iii
<b>ACKNOWLEDGEMENTS</b>	v
<b>LIST OF FIGURES</b>	xii
<b>LIST OF TABLES</b>	xv
<b>CHAPTER 1 INTRODUCTION</b>	1
<b>CHAPTER 2 ADVANCES IN GRADIENT-BASED OPTIMIZATION</b>	7
2.1 Introduction	7
2.2 Formulations of $\ell_p$ Optimization	9
2.2.1 $\ell_p$ norms, One-Sided and Generalized $\ell_p$ Functions	9
2.2.2 Nonlinear Programming Equivalent Problems	11
2.2.3 Linear $\ell_p$ Optimization	12
2.3 Trust Region Gauss-Newton Method	14
2.3.1 Semi-Linearization of a Nonlinear $\ell_p$ Function	14
2.3.2 Trust Regions	15
2.3.3 The Levenberg-Marquardt Method	17
2.4 Quasi-Newton Method Applied to an $\ell_p$ Problem	17
2.4.1 Solution of the Optimality Equations	18
2.4.2 Updates of a Hessian	19
2.5 Combined Methods	21
2.6 An Algorithm for Linearly Constrained One-Sided $\ell_1$ Optimization	22
2.6.1 Introductory Remarks	22
2.6.2 Stage 1: A Trust Region Gauss-Newton Method	23
2.6.3 Stage 2: A Quasi-Newton Method	25
2.6.4 A Combined 2-Stage Algorithm	28

**TABLE OF CONTENTS (continued)**

	PAGE
<b>CHAPTER 2    ADVANCES IN GRADIENT-BASED OPTIMIZATION</b> (continued)	
2.7 A Unified Fortran Library for Nonlinear Optimization	29
2.8 Concluding Remarks	30
<b>CHAPTER 3    EFFICIENT APPROACHES TO NETWORK SENSITIVITY</b> <b>              ANALYSIS</b>	32
3.1 Introduction	32
3.2 Sensitivity Analysis Using a Nodal Description	33
3.3 General M-Ports	34
3.3.1 Unterminated M-Ports	34
3.3.2 Terminated M-Ports	37
3.4 Two-Ports	38
3.4.1 General Two-Ports	38
3.4.2 Second-Order Sensitivities	39
3.4.3 Sensitivities of Commonly Used Frequency Responses	40
3.4.4 Lossless Two-Ports	40
3.5 Concluding Remarks	46
<b>CHAPTER 4    NOMINAL CIRCUIT OPTIMIZATION</b>	48
4.1 Introduction	48
4.2 Basic Formulation	49
4.2.1 A Hierarchy of Simulation Models	49
4.2.2 Specifications and Error Functions	53
4.2.3 Variables and Objective Functions	57
4.3 Optimization of Multi-Coupled Cavity Filters	58
4.3.1 The Physical Structure and the Equivalent Circuit	59
4.3.2 Efficient Simulation and Sensitivity Calculations	61
4.3.3 Cubic Interpolation in Minimax Design	65
4.3.4 10th-Order Elliptic and Quasi-Elliptic Self-Equalized Filters	66
4.3.5 An Asymmetric Design	70

**TABLE OF CONTENTS (continued)**

	PAGE
<b>CHAPTER 4 NOMINAL CIRCUIT OPTIMIZATION (continued)</b>	
4.3.6 First-Order Prediction of the Effect of Cavity Dissipation	74
4.4 Large-Scale Optimization of Manifold Multiplexers	77
4.5 Concluding Remarks	82
<b>CHAPTER 5 REALISTIC APPROACHES TO CIRCUIT DESIGN</b>	<b>85</b>
5.1 Introduction	85
5.2 A Multi-Circuit Formulation	86
5.2.1 Physical Tolerances and Model Uncertainties	86
5.2.3 Multiple Circuits and Yield	88
5.2.4 Centering, Tolerancing and Tuning	91
5.3 A Review of Techniques for Statistical Design	92
5.3.1 Worst-Case Design	97
5.3.2 Approximations of the Acceptable Region	98
5.3.3 The Gravity Method	100
5.3.4 The Parametric Sampling Method	100
5.4 A Generalized $\ell_p$ Centering Algorithm	102
5.4.1 Representing an Outcome by an $\ell_p$ Function	102
5.4.2 Generalized $\ell_p$ Centering	104
5.4.3 Tolerancing and Tuning	107
5.4.4 Circuit Examples	109
5.5 Concluding Remarks	114
<b>CHAPTER 6 A MULTI-CIRCUIT APPROACH TO DEVICE MODELING</b>	<b>116</b>
6.1 Introduction	116
6.2 Extension of the Nominal Design Concept to Modeling	117
6.3 Modeling Using Multiple Circuits	118
6.3.1 Uncertainties that Affect Modeling	118
6.3.2 Multiple Circuits and Common Variables	120
6.3.3 Computational Considerations	123
6.3.4 Exploiting the Unique Property of the $\ell_1$ Optimization	125

**TABLE OF CONTENTS (continued)**

	PAGE
<b>CHAPTER 6 A MULTI-CIRCUIT APPROACH TO DEVICE MODELING (continued)</b>	
6.4 Practical Applications	126
6.4.1 Unique Identification of Model Parameters	126
6.4.2 Multi-Circuit Modeling of a FET Device	132
6.4.3 Model Verification for a Multi-Cavity Filter	135
6.4.4 Modeling of the Relationship Between Physical and Equivalent Circuit Parameters	142
6.5 Concluding Remarks	146
<b>CHAPTER 7 OPTIMIZATION WITH INTEGRATED GRADIENT APPROXIMATIONS</b>	150
7.1 Introduction	150
7.2 Gradient Approximations	152
7.2.1 Estimating Gradient by Perturbations	152
7.2.2 The Broyden Update	153
7.2.3 A Weighted Broyden Update	155
7.2.4 The Special Iterations of Powell	157
7.2.5 A Hybrid Approximation Algorithm	157
7.2.6 Integration with Optimization Methods	159
7.3 A Minimax Algorithm Using Approximate Gradients	160
7.3.1 Implementation of a Fortran Subroutine	160
7.3.2 Performance on Some Test Problems	161
7.3.3 Worst-Case Design of a Microwave Amplifier	169
7.3.4 Design of a 5-Channel Multiplexer	171
7.4 An $\ell_1$ Algorithm Using Approximate Gradients	187
7.4.1 Implementation of a Fortran Subroutine	187
7.4.2 Performance on Some Test Problems	188
7.4.3 Fault Location of a Mesh Network	194
7.4.4 Multi-Circuit Modeling of a FET Device	196
7.5 Concluding Remarks	198
<b>CHAPTER 8 CONCLUSIONS</b>	200
<b>APPENDIX FORMULAS FOR POWELL'S SPECIAL ITERATIONS</b>	205
<b>REFERENCES</b>	206

**TABLE OF CONTENTS (continued)**

	PAGE
<b>AUTHOR INDEX</b>	214
<b>SUBJECT INDEX</b>	218

## LIST OF FIGURES

FIGURE		PAGE
3.1	An illustration of multi-ports.	36
4.1	A microwave stripline transformer.	52
4.2	Illustrations of upper and lower specifications, error functions and the generalized $\ell_p$ objective functions.	55
4.3	Illustrations of single specifications; error functions and the $\ell_p$ norms.	56
4.4	Typical structures of multi-coupled cavity filters.	60
4.5	Equivalent circuit for an unterminated multi-coupled cavity filter.	62
4.6	Responses of the optimized 10th-order elliptic multi-coupled cavity filter.	69
4.7	Responses of the optimized 10th-order quasi-elliptic self-equalized filter.	72
4.8	The poles and zeros of the transfer function of the 10th-order self-equalized filter.	73
4.9	Responses of the asymmetric 6th-order filter.	75
4.10	Passband insertion loss of the 10th-order self-equalized filter for $Q = 10,000$ and $Q = \infty$ .	78
4.11	A typical multiplexer structure.	79
4.12	Responses of the optimized 12 channel multiplexer.	81
4.13	Optimized responses of the 16 channel multiplexer.	83
5.1	Three nominal points and the related yield.	90
5.2	Illustrations of tuning.	93
5.3	Illustration of multi-circuit design considering eight outcomes.	94
5.4	A typical cost-versus-yield curve.	95

**LIST OF FIGURES (continued)**

FIGURE		PAGE
5.5	A maximum yield design and a minimum cost design.	96
5.6	A Chebyshev lowpass filter.	110
6.1	Illustration of common and independent variables in multi-circuit modeling.	122
6.2	Approximations using $\ell_1$ and $\ell_2$ optimization.	127
6.3	A simple RC network.	130
6.4	Small-signal equivalent circuit model for a FET device.	133
6.5	FET measurements and model responses for the biasing conditions $V_{ds} = 4V$ , $V_{gs} = 0V$ and $I_{ds} = 177mA$ .	137
6.6	FET measurements and model responses for the biasing conditions $V_{ds} = 4V$ , $V_{gs} = -1.74V$ and $I_{ds} = 92mA$ .	138
6.7	FET measurements and model responses for the biasing conditions $V_{ds} = 4V$ , $V_{gs} = -3.1V$ and $I_{ds} = 37mA$ .	139
6.8	Measured and modeled responses of the 8th-order filter before adjusting the iris.	143
6.9	Measured and modeled responses of the 8th-order filter after adjusting the iris.	144
6.10	The identified coupling values of a 6th-order filter versus the relative position of the screw which controls $M_{12}$ dominantly.	147
6.11	The identified coupling values of a 6th-order filter versus the relative position of the screw which controls $M_{34}$ dominantly.	148
6.12	The identified coupling values of a 6th-order filter versus the relative position of the screw which controls $M_{56}$ dominantly.	149
7.1	Two-section, 10:1 transmission line transformer.	163
7.2	Minimax solution of problem M1 by exact gradients.	164
7.3	Minimax solution of problem M1 using approximate gradients.	166

LIST OF FIGURES (continued)

FIGURE		PAGE
7.4	Minimax solution of problem M2 by exact gradients.	167
7.5	Minimax solution of problem M2 using approximate gradients.	168
7.6	A microwave amplifier.	170
7.7	Worst-case envelope for the amplifier response at the centered solution.	173
7.8	Responses of the 5-channel multiplexer at the starting point.	176
7.9	Responses of the 5-channel multiplexer at the solution obtained using exact gradients.	177
7.10	Responses of the 5-channel multiplexer at the solution obtained using only the Broyden update and special iterations for gradient approximation.	181
7.11	Responses of the 5-channel multiplexer obtained after 500 response evaluations using approximate gradients with regular corrections.	183
7.12	Responses of the 5-channel multiplexer at the solution using approximate gradients with regular corrections.	184
7.13	Responses of the 5-channel multiplexer at the solution using the weighted update.	186
7.14	The $\ell_1$ solution of problem L1 using perturbations for the gradients.	191
7.15	The $\ell_1$ solution of problem L1 using approximate gradients.	192
7.16	A resistive mesh network.	195

## LIST OF TABLES

TABLE		PAGE
2.1	Nonlinear programming equivalent formulations for the $\ell_p$ optimization.	13
3.1	Sensitivity expressions for selected frequency responses.	41
4.1	Parameters for the 10th-order elliptic filter.	68
4.2	Parameters for the 10th-order quasi-elliptic self-equalized filter.	71
4.3	Parameters for the 6th-order asymmetric filter.	76
5.1	Generalized $\ell_1$ centering of the Chebyshev lowpass filter.	112
5.2	Generalized $\ell_1$ centering of a multi-coupled cavity filter.	113
6.1	Parameter values of the FET models.	136
6.2	Identified parameter values for the 8th-order filter.	141
7.1	Comparison of computational effort for the minimax examples.	162
7.2	Parameter values of the microwave amplifiers.	172
7.3	Multiplexer center frequencies and bandwidths.	178
7.4	Multiplexer parameters optimized using exact gradients.	180
7.5	Multiplexer parameters optimized using approximate gradients.	185
7.6	Comparison of computational effort for examples L1 to L4.	189
7.7	Comparison of computational effort for example L5.	190



## CHAPTER 1

### INTRODUCTION

The continuing effort to formulate and solve increasingly complex engineering problems through the state-of-the-art mathematical optimization represents one of the driving forces of advanced study in computer-aided design (CAD). The astonishing progress in computer hardware, leading to drastic reduction in the cost of mass computation and the widespread use of personal computers, has given further impetus to the development of efficient CAD techniques.

In electrical engineering, one of the earliest applications of CAD techniques is in the area of filter design. Methods that were popular at the time have been summarized in the classic paper by Temes and Calahan (1967). Since then, advances have been made in many directions. Optimization techniques have evolved from simple and low-dimension-oriented methods into sophisticated and powerful tools serving practicing engineers. Efficient approaches have been developed for large-scale network simulation and sensitivity calculations.

In recent years, there has been a continuing trend toward dealing more explicitly with process imprecision, manufacturing tolerances, model uncertainties, measurement errors, and so on. Such realistic considerations arise from design problems in which a large volume of production is envisaged, e.g., integrated circuits. They also arise from modeling problems in which consistent results

are expected despite measurement limitations, model approximations and simplifications.

This thesis offers a unified and integrated treatment of three essential aspects of circuit CAD: effective use of the state-of-the-art optimization tools, efficient calculation of exact and approximate gradients, and adequate mathematical representation of engineering design and modeling problems. Emphasis is given to recent, but well-tested, advances in gradient-based minimax,  $\ell_1$  and  $\ell_2$  optimization. To provide the required gradients, elegant and efficient approaches to network sensitivity analysis are described. Since the difficulties in exact sensitivity calculation for some applications have contributed a sizable gap between the advanced optimization theories and their actual implementation, we also develop an efficient and integrated algorithm for gradient approximations. A hierarchical representation is identified for the nominal circuit which is then extended to including tolerances and model uncertainties, resulting in a consistent formulation of the  $\ell_p$  circuit optimization. A multi-circuit approach is presented which unifies the concepts of realistic circuit design and robust device modeling.

The methods described for design, modeling and gradient approximations are applied to relevant practical examples such as multi-coupled cavity filters, waveguide manifold multiplexers and FET devices. Not only serving to illustrate the theory, these examples are also of current significance to researchers and engineers, especially in the field of satellite communications.

Chapter 2 is concerned with the recent theoretical and algorithmic advances in the  $\ell_p$  optimization. The definitions and properties of the  $\ell_p$  norms, one-sided and generalized  $\ell_p$  functions are reviewed. Following the essence of the Hald and Madsen (1981, 1985) algorithms for minimax and  $\ell_1$ , we describe the trust region Gauss-Newton method, which solves a sequence of semi-linearized  $\ell_p$  problems, and the quasi-Newton method which is applied to solving the optimality equations. We then present a new algorithm for linearly constrained one-sided  $\ell_1$  optimization which is very useful in circuit design and centering. Based on the same principle as the Hald and Madsen approach, our algorithm is a 2-stage combined method. Linear programming techniques are utilized to find a trust region solution to a linearized subproblem. The optimality equations for the one-sided  $\ell_1$  problem are also derived to which a quasi-Newton iteration is applied.

Chapter 3 deals with efficient and systematic calculation of network sensitivities. A simplified algebraic approach to linear network sensitivity analysis is reviewed. Useful formulas are derived for unterminated and terminated general multi-ports. The results are specialized to two-ports which are widely used to represent filters and subnetworks. Elegant and original proofs are shown for an important formula for lossless reciprocal two-ports. Sensitivity expressions for some commonly used frequency responses are also given.

Chapter 4 is devoted to nominal circuit optimization. A hierarchy of simulation models is introduced, and the parameters

and responses associated with these models are identified. Error functions and an  $\ell_p$  objective are formulated from the model responses and the performance specifications. The nominal minimax optimization of multi-coupled cavity filters, whose large variety and complexity have made them prime candidates for CAD, is described in detail. Three examples of practical interest, including elliptic, self-equalized and asymmetric filters, are presented. Large-scale circuit optimization is demonstrated by a 16-channel multiplexer design.

Chapter 5 considers realistic circuit design. Tolerances and uncertainties associated with the models of different levels are exposed. Multiple circuits are defined to relate these uncertainties to a nominal point. The concepts of design centering, tolerancing and tuning, with the aim of improving the yield and reducing the production cost, are discussed. Techniques for statistical design are reviewed and several representative methods are shown in some detail. A generalized  $\ell_p$  centering algorithm is proposed which offers a natural extension to nominal  $\ell_p$  optimization and a unified approach to yield enhancement.

In Chapter 6, we examine the motivation, theoretical foundation and practical applications of a novel approach to device modeling. Attempting to overcome the adverse effects of various uncertainties to modeling, the new approach utilizes the concept of simultaneous processing of multiple circuits and exploits the theoretical properties of the  $\ell_1$  optimization. Unlike the traditional approach, it emphasizes the uniqueness and consistency of the res-

ults of modeling. Applications to parameter identification and model verification are illustrated by industrial examples.

Chapter 7 addresses itself to the subject of optimization with integrated gradient approximations. An algorithm combining perturbations, the Broyden update (Broyden 1965) and the special iterations (Powell 1970) is described. A weighted formula which may improve the performance of the Broyden update is developed. Integration of the gradient approximations with an optimization routine is discussed and illustrated through the minimax and  $\ell_1$  implementations. The effectiveness and efficiency of the proposed approach are demonstrated by abundant examples.

We conclude in Chapter 8 with some suggestions for further research.

The author contributed substantially to the following original developments presented in this thesis:

- (1) A new algorithm for the linearly constrained one-sided  $\ell_1$  optimization.
- (2) New proofs of a sensitivity formula for reciprocal lossless two-ports.
- (3) An efficient approach to the simulation, sensitivity analysis and optimization of multi-coupled cavity filters.
- (4) A generalized  $\ell_p$  centering algorithm for statistical design centering and yield enhancement.
- (5) Theoretical results of a novel approach to device modeling which utilizes multiple circuits and its application to FET modeling.

- (6) A general approach to gradient approximations and the use of a weighted update.
- (7) Integration of the gradient approximation method with the minimax and  $\ell_1$  algorithms and its practical applications.

## CHAPTER 2

### ADVANCES IN GRADIENT-BASED OPTIMIZATION

#### 2.1 INTRODUCTION

In the last two decades, the advances in the theoretical and algorithmic aspects of optimization techniques have been truly astonishing. Nonlinear optimization has become not only a subject for academic research, but also a powerful tool serving practicing engineers. Modern state-of-the-art methods have largely replaced the primitive trial-and-error approach. Especially, gradient-based optimization methods have gained increasing popularity in recent years, since they demonstrate in general a far superior performance to direct (non-gradient) methods.

Traditionally, the minimization of a least squares measure has been favored for its relative simplicity and differentiability. However, increasingly complex and diversified engineering applications have demanded more sophisticated techniques, e.g., the use of minimax and  $\ell_1$  norms and multiple objectives (multiple criteria).

This chapter concerns itself with the recent advances in  $\ell_p$  (least pth) optimization. The earliest circuit applications of the least pth approximation were addressed by Temes and Zai (1969). Later, Bandler and Charalambous (1972) have developed the theory of generalized  $\ell_p$  optimization. Recently, the approach due to Hald and Madsen (1981, 1985, also Madsen 1985) has proved highly successful in solving minimax and  $\ell_1$  problems.

We begin this chapter with the formulation of a least pth measure  $H(\mathbf{f})$  which may be an  $l_p$  norm, a one-sided  $l_p$  function or a generalized  $l_p$  function. In circuit optimization,  $\mathbf{f}$  typically represents a set of discretized error functions with respect to the specified (in design) or measured (in modeling) circuit responses. In this chapter we concentrate on the methods of minimizing  $H(\mathbf{f})$ , leaving the construction of  $\mathbf{f}$  to subsequent chapters.

Following the essence of the Hald and Madsen approach, we describe the trust region Gauss-Newton method and the quasi-Newton method applied to solving optimality equations. We establish the nonlinear programming equivalent to  $l_p$  optimization which leads to the solutions of linear minimax,  $l_2$  and  $l_1$  problems by linear or quadratic programming. In the Gauss-Newton method, a sequence of semi-linearized  $l_p$  problems are defined and a trust region solution is found for each problem. This has been shown (Madsen 1985) to provide global convergence of the algorithm. From the Kuhn-Tucker (1951) conditions for nonlinear programming we can derive a set of optimality equations for an  $l_p$  problem and solve these equations by a quasi-Newton method. Hald and Madsen (1981, 1985) have combined the Gauss-Newton and quasi-Newton methods into reliable and powerful algorithms for minimax and  $l_1$  optimization.

We then present a new algorithm for linearly constrained one-sided  $l_1$  optimization based on the Hald and Madsen approach. It is a 2-stage algorithm combining a trust region Gauss-Newton method and a quasi-Newton iteration.

Finally, a unified Fortran library for minimax,  $l_2$ ,  $l_1$  and

one-sided  $\ell_1$  optimization is described.

## 2.2 FORMULATIONS OF $\ell_p$ OPTIMIZATION

### 2.2.1 $\ell_p$ Norms, One-Sided and Generalized $\ell_p$ Functions

Given a set of nonlinear functions

$$\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}) \ f_2(\mathbf{x}) \ \dots \ f_m(\mathbf{x})]^T, \quad (2.1)$$

where the superscript T stands for vector or matrix transposition, and

$$\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T \quad (2.2)$$

is the set of variables, the  $\ell_p$  norm of  $\mathbf{f}$  is defined as (Temes and Zai 1969)

$$\|\mathbf{f}\|_p = \left[ \sum_{j=1}^m |f_j|^p \right]^{1/p}. \quad (2.3)$$

Least squares ( $\ell_2$ ) is perhaps the most widely used norm, which is given by

$$\|\mathbf{f}\|_2 = \left[ \sum_{j=1}^m |f_j|^2 \right]^{1/2}. \quad (2.4)$$

In practice, we often use  $\|\mathbf{f}\|_2^2$  as an objective function which is differentiable and its gradient can be easily obtained from the partial derivatives of  $\mathbf{f}$ . Furthermore, it is a quadratic function if  $\mathbf{f}$  is linear in  $\mathbf{x}$ . Partly due to these properties, a large variety of  $\ell_2$  optimization techniques have been developed and popularly implemented in CAD software (e.g., TOUCHSTONE 1985 and SUPER-COMPACT 1986).

The parameter  $p$  has an important implication. By choosing a large (small) value for  $p$ , we in effect place more emphasis on

those error functions ( $f_j$ 's) that have larger (smaller) values. Letting  $p = \infty$ , we have the minimax norm

$$\|f\|_{\infty} = \max_j |f_j| \quad (2.5)$$

which directs all the attention to the worst case and the other errors are in effect ignored. The minimax norm is suitable if we wish to minimize  $f$  in an optimal equal-ripple manner (the worst  $f_j$ 's being equal in magnitude), for example, in filter design.

On the other hand, the use of the  $\ell_1$  norm, as defined by

$$\|f\|_1 = \sum_{j=1}^m |f_j|, \quad (2.6)$$

implies attaching more importance to the error functions that are closer to zero. This property is often exploited in data-fitting problems (e.g., Bartels and Conn 1981). The application of the  $\ell_1$  norm to modeling will be exposed in Chapter 6.

Notice that neither  $\|f\|_{\infty}$  nor  $\|f\|_1$  is differentiable in the ordinary sense. Therefore, their minimization requires algorithms that are much more sophisticated than those in the  $\ell_2$  case.

With the  $\ell_p$  norm, we are minimizing the error functions towards a zero value. Suppose that our true intention is to have  $f_j \leq 0$ , then a negative value of  $f_j$  simply indicates that the goal is exceeded and is, in a sense, better than having  $f_j = 0$ . This fact leads to the use of the one-sided  $\ell_p$  function defined by

$$H_p^+(f) = \left[ \sum_{j \in J} |f_j|^p \right]^{1/p}, \quad (2.7)$$

where  $J = \{j \mid f_j \geq 0\}$ . Actually, if we define  $f_j^+ = \max\{f_j, 0\}$ , then  $H_p^+(f) = \|f^+\|_p$ .

Bandler and Charalambous (1972) have proposed the use of a generalized  $\ell_p$  function defined by

$$H_p(\mathbf{f}) = \begin{cases} H_p^+(\mathbf{f}) & \text{if the set } J \text{ is not empty} \\ H_p^-(\mathbf{f}) & \text{otherwise} \end{cases}, \quad (2.8)$$

where

$$H_p^-(\mathbf{f}) = -\left[ \sum_{j=1}^m (-f_j)^{-p} \right]^{-1/p}. \quad (2.9)$$

In other words, when at least one  $f_j$  is nonnegative we use  $H_p^+$ , and  $H_p^-$  is defined if all the error functions have become negative.

Compared to the one-sided function, the generalized  $\ell_p$  function has an advantage in that it is meaningfully defined for the case where all the  $f_j$  are negative, thus permitting further minimization of the objective function.

A classical example with extensive circuit applications is the generalized minimax function

$$H_\infty(\mathbf{f}) = \max_j \{f_j\}, \quad (2.10)$$

which can be found in, for instance, the design of Chebyshev type bandpass filters.

### 2.2.2 Nonlinear Programming Equivalent Problems

$\ell_1$ ,  $\ell_2$  and  $\ell_\infty$  are by far the most useful members of the  $\ell_p$  family. Apart from their distinctive theoretical properties, it is very important from the algorithmic point of view that the exact solutions to linear  $\ell_1$ ,  $\ell_2$  and  $\ell_\infty$  problems can be found by linear or quadratic programming. Besides, the other members of the  $\ell_p$  family have a continuously differentiable objective function and,

therefore, can be treated similarly to the  $\ell_2$  case.

In this section, we define nonlinear programs which are equivalent to  $\ell_1$ ,  $\ell_2$  and  $\ell_\infty$  problems. The equivalent formulations often clarify the concepts of local linearization and optimality conditions.

For instance, the minimization of  $\|\mathbf{f}\|_1$  is equivalent to

$$\begin{aligned} & \text{minimize}_{\mathbf{x}, \mathbf{y}} \sum_{j=1}^m y_j \\ & \text{subject to} \end{aligned} \tag{2.11}$$

$$y_j \geq f_j(\mathbf{x}), \quad y_j \geq -f_j(\mathbf{x}), \quad j = 1, 2, \dots, m.$$

The one-sided  $\ell_1$  problem can be treated as

$$\begin{aligned} & \text{minimize}_{\mathbf{x}, \mathbf{y}} \sum_{j=1}^m y_j \\ & \text{subject to} \end{aligned} \tag{2.12}$$

$$y_j \geq f_j(\mathbf{x}), \quad y_j \geq 0, \quad j = 1, 2, \dots, m.$$

These results, as well as those for the minimax and least squares, are summarized in Table 2.1. For the convenience of presentation, we denote these nonlinear programs by  $P(\mathbf{x}, \mathbf{f})$ .

### 2.2.3 Linear $\ell_p$ Optimization

A linear  $\ell_p$  problem implies that the set of error functions  $\mathbf{f}$  is linear in the variables  $\mathbf{x}$ . As an important consequence, the equivalent problem  $P(\mathbf{x}, \mathbf{f})$  becomes a linear or quadratic program whose exact solution can be found using standard techniques (e.g., the simplex method for linear programming due to Dantzig 1951).

TABLE 2.1  
 NONLINEAR PROGRAMMING EQUIVALENT FORMULATIONS  
 FOR  $l_1$ ,  $l_2$  AND  $l_\infty$  OPTIMIZATION

---

Original problem:	minimize $H(\mathbf{f})$ $\mathbf{x}$	
Equivalent problem:	minimize $V(\mathbf{x}, \mathbf{y})$ subject to the constraints $\mathbf{x}, \mathbf{y}$	

---

H(f)	V(x,y)	constraints (for $j = 1, 2, \dots, m$ )
$\ \mathbf{f}\ _1$	$\sum_{j=1}^m y_j$	$y_j \geq f_j, \quad y_j \geq -f_j$
$\ \mathbf{f}\ _2$	$\mathbf{y}^T \mathbf{y}$	$y_j = f_j$
$\ \mathbf{f}\ _\infty$	$y$	$y \geq f_j, \quad y \geq -f_j$
$H_1^+(\mathbf{f})$	$\sum_{j=1}^m y_j$	$y_j \geq f_j, \quad y_j \geq 0$
$H_2^+(\mathbf{f})$	$\mathbf{y}^T \mathbf{y}$	$y_j \geq f_j, \quad y_j \geq 0$
$H_\infty^+(\mathbf{f})$	$y$	$y \geq f_j, \quad y \geq 0$
$H_\infty(\mathbf{f})$	$y$	$y \geq f_j$

---

For instance, consider a linear  $\ell_1$  problem. Let the set of linear functions be

$$f_j = \mathbf{c}_j^T \mathbf{x} + d_j, \quad j = 1, 2, \dots, m, \quad (2.13)$$

where  $\mathbf{c}_j$  and  $d_j$  are constants. Then the solution to the following linear program

$$\begin{aligned} & \underset{\mathbf{x}, \mathbf{y}}{\text{minimize}} && \sum_{j=1}^m y_j \\ & \text{subject to} && \\ & y_j - \mathbf{c}_j^T \mathbf{x} - d_j \geq 0, && j = 1, 2, \dots, m, \\ & y_j + \mathbf{c}_j^T \mathbf{x} + d_j \geq 0, && \end{aligned} \quad (2.14)$$

is the minimizer of  $\|\mathbf{f}\|_1$ .

An equally important fact is that linear constraints can be easily incorporated into the solution of a linear  $\ell_p$  problem. Let  $P(\mathbf{x}, \mathbf{f}, D)$  be the problem of  $P(\mathbf{x}, \mathbf{f})$  subject to a set of linear constraints of the form

$$D: \begin{aligned} \mathbf{a}_k^T \mathbf{x} + b_k &= 0, & k &= 1, 2, \dots, L_{\text{eq}}, \\ \mathbf{a}_k^T \mathbf{x} + b_k &\geq 0, & k &= L_{\text{eq}}+1, \dots, L, \end{aligned} \quad (2.15)$$

where  $\mathbf{a}_k$  and  $b_k$  are constants. If  $P(\mathbf{x}, \mathbf{f})$  is a linear or quadratic program, so is  $P(\mathbf{x}, \mathbf{f}, D)$ . In other words, linearly constrained linear  $\ell_1$ ,  $\ell_2$  and  $\ell_\infty$  problems can also be solved using standard linear or quadratic programming techniques.

## 2.3 TRUST REGION GAUSS-NEWTON METHOD

### 2.3.1 Semi-Linearization of a Nonlinear $\ell_p$ Function

For a general problem  $P(\mathbf{x}, \mathbf{f})$ , if we substitute  $\mathbf{f}$  with a set of linearized functions  $\bar{\mathbf{f}}$ , then  $P(\mathbf{x}, \bar{\mathbf{f}})$  becomes a linear  $\ell_p$  problem

which can be solved using standard techniques. In a Gauss-Newton method, we define and solve a sequence of such linear subproblems.

At each iteration, given  $\mathbf{x}_k$ , a set of linearized functions is defined as

$$\bar{\mathbf{f}}(\mathbf{h}) = \mathbf{f}(\mathbf{x}_k) + \mathbf{G}(\mathbf{x}_k)\mathbf{h}, \quad (2.16)$$

where  $\mathbf{G}$  is the Jacobian given by

$$\mathbf{G}(\mathbf{x}) = [\partial \mathbf{f}^T / \partial \mathbf{x}]^T. \quad (2.17)$$

We then solve the linear subproblem  $P(\mathbf{h}, \bar{\mathbf{f}})$  using linear or quadratic programming. From a slightly different viewpoint, we may call this a semi-linearization (Madsen 1985) of the nonlinear objective function  $U(\mathbf{x}) = H(\mathbf{f})$  resulting in

$$\bar{U}(\mathbf{h}) = H(\bar{\mathbf{f}}(\mathbf{h})). \quad (2.18)$$

It should be noted that (2.18) is quite different from a normal linearization as  $U(\mathbf{h}) \approx U(\mathbf{x}_k) + [\mathbf{U}'(\mathbf{x}_k)]^T \mathbf{h}$  which corresponds to a steepest descent method. In fact the gradient  $\mathbf{U}'$  may not even exist.

### 2.3.2 Trust Regions

In general, the set of linearized functions  $\bar{\mathbf{f}}$  as given by (2.16) is valid only as a local model and, therefore, its use should be restricted to within a neighbourhood of  $\mathbf{x}_k$  in which  $\bar{\mathbf{f}}$  is believed to be a good approximation to  $\mathbf{f}$ . Such a neighbourhood is called, appropriately, a trust region and typically defined by

$$D: \begin{aligned} \Lambda_k &\geq h_j, & i = 1, 2, \dots, n, \\ \Lambda_k &\geq -h_j, & i = 1, 2, \dots, n. \end{aligned} \quad (2.19)$$

At each iteration, we incorporate a trust region and solve

a linearly constrained linear  $\ell_p$  subproblem  $P(\mathbf{h}, \bar{\mathbf{f}}, D)$ . Denote its solution by  $\mathbf{h}_k$ . If  $\mathbf{x}_k + \mathbf{h}_k$  reduces the original nonlinear objective function we take it as the next iterate, i.e., if  $U(\mathbf{x}_k + \mathbf{h}_k) < U(\mathbf{x}_k)$  then  $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{h}_k$ , otherwise we let  $\mathbf{x}_{k+1} = \mathbf{x}_k$ . In the latter case, the trust region is apparently too large and, consequently, should be reduced.

Hald and Madsen (1981, 1985) suggested that the local bound  $\Lambda_k$  be adjusted according to the goodness of the linearized model. More precisely, if

$$U(\mathbf{x}_k) - U(\mathbf{x}_k + \mathbf{h}_k) \leq \delta_1 [U(\mathbf{x}_k) - \bar{U}(\mathbf{h}_k)], \quad (2.20)$$

then the trust region appears to be too large and the bound is decreased:  $\Lambda_{k+1} = K_1 \Lambda_k$ . Otherwise, if

$$U(\mathbf{x}_k) - U(\mathbf{x}_k + \mathbf{h}_k) \geq \delta_2 [U(\mathbf{x}_k) - \bar{U}(\mathbf{h}_k)], \quad (2.21)$$

then the bound is increased:  $\Lambda_{k+1} = K_2 \Lambda_k$ . If neither (2.20) nor (2.21) holds then  $\Lambda_{k+1} = \Lambda_k$ . The constants  $\{\delta_1, \delta_2, K_1, K_2\}$  should satisfy  $0 < \delta_1 < \delta_2 < 1$  and  $0 < K_1 < 1 < K_2$ .

Madsen (1985) has shown that the trust region Gauss-Newton method provides global convergence in which the proper use of trust regions constitutes a critical part. In some other earlier work by Osborne and Watson (1969, 1971) the problem  $P(\mathbf{h}, \bar{\mathbf{f}})$  was solved without incorporating a trust region and the solution  $\mathbf{h}_k$  was then used as the direction for a line search. For their algorithms no convergence can be guaranteed and  $\{\mathbf{x}_k\}$  may even converge to a non-stationary point.

The trust region methods in a broader context have been surveyed recently by Moré (1982).

### 2.3.3 The Levenberg-Marquardt Method

Normally for the least squares problem we have to solve a quadratic program at each iteration, which can be a time-consuming process. A remarkable alternative is the method due to Levenberg (1944) and Marquardt (1963). Given  $\mathbf{x}_k$ , it defines

$$\underset{\mathbf{h}}{\text{minimize}} \mathbf{h}^T (\mathbf{G}^T \mathbf{G} + \theta_k \mathbf{1}) \mathbf{h} + 2\mathbf{f}^T \mathbf{G} \mathbf{h} + \mathbf{f}^T \mathbf{f} , \quad (2.22)$$

where  $\mathbf{G} = \mathbf{G}(\mathbf{x}_k)$ ,  $\mathbf{f} = \mathbf{f}(\mathbf{x}_k)$  and  $\mathbf{1}$  is an identity matrix. The minimizer  $\mathbf{h}_k$  is obtained simply by solving a linear system

$$(\mathbf{G}^T \mathbf{G} + \theta_k \mathbf{1}) \mathbf{h}_k = \mathbf{G}^T \mathbf{f} \quad (2.23)$$

using, for example, LU factorization.

The Levenberg-Marquardt parameter  $\theta_k$  is very critical for this method. First of all, it is chosen to guarantee the positive definiteness of (2.23). Furthermore, it plays, roughly speaking, an inverse role of  $\Lambda_k$  to control the size of a trust region. When  $\theta_k \rightarrow \infty$ ,  $\mathbf{h}_k$  gives an infinitesimal steepest descent step. When  $\theta_k = 0$ ,  $\mathbf{h}_k$  becomes the solution to  $P(\mathbf{h}, \mathbf{f})$  without bounds, which is equivalent to having  $\Lambda_k \rightarrow \infty$ . Hence, the rules for updating  $\theta_k$  should be opposite to those for  $\Lambda_k$ .

## 2.4 QUASI-NEWTON METHOD APPLIED TO AN $\ell_p$ PROBLEM

Quasi-Newton methods (variable metric methods) emerged from the original work of Davidon (1959), Fletcher and Powell (1963), as well as Broyden (1965, 1967).

For a differentiable objective function  $U$ , a quasi-Newton step is given by

$$\mathbf{h}_k = - \alpha_k \mathbf{B}_k^{-1} \mathbf{U}'(\mathbf{x}_k), \quad (2.24)$$

where  $\mathbf{B}_k$  is an approximation to the Hessian of  $U(\mathbf{x})$  and the step size controlling parameter  $\alpha_k$  is to be determined through a line search. However, for the  $\ell_1$  and the minimax objective functions, the gradient  $\mathbf{U}'$  may not exist, much less the Hessian.

#### 2.4.1 Solution of the Optimality Equations

The optimality conditions for an  $\ell_p$  problem provide more insight to the general case. Applying the Kuhn-Tucker conditions (Kuhn and Tucker 1951) to the nonlinear programming problem  $P(\mathbf{x}, \mathbf{f})$ , we shall find a set of optimality equations

$$\mathbf{R}(\mathbf{x}) = \mathbf{0} \quad (2.25)$$

which must be satisfied by a local optimum  $\mathbf{x}^*$ . Naturally, we are motivated to solve (2.25), as a means of finding the minimizer of  $U(\mathbf{x})$ . A quasi-Newton step for solving nonlinear equations (2.25) is given by

$$\mathbf{h}_k = -\alpha_k \mathbf{J}_k^{-1} \mathbf{R}(\mathbf{x}_k), \quad (2.26)$$

where  $\mathbf{J}_k$  is an approximate Jacobian of  $\mathbf{R}(\mathbf{x})$ . Only when  $U(\mathbf{x})$  is differentiable will the optimality equations be  $\mathbf{R}(\mathbf{x}) = \mathbf{U}'(\mathbf{x}) = \mathbf{0}$  and (2.26) reverts to the more familiar formula (2.24).

Consider the  $\ell_1$  case as an example. The optimality equations can be shown as (Charalambous 1979)

$$\mathbf{R}(\mathbf{x}, \mathbf{g}) = \begin{bmatrix} \sum_{j \notin Z} \mu_j \mathbf{f}'_j(\mathbf{x}) + \sum_{j \in Z} \delta_j \mathbf{f}'_j(\mathbf{x}) \\ \mathbf{f}_Z(\mathbf{x}) \end{bmatrix} = \mathbf{0}, \quad (2.27)$$

where  $Z$  is an index set, as  $Z(\mathbf{x}^*) = \{j \mid \mathbf{f}_j(\mathbf{x}^*) = 0\}$ , identifying the zero functions at the optimum. The vector  $\mathbf{f}_Z(\mathbf{x})$  consists of the functions in the set  $Z$  and the multipliers satisfy  $|\delta_j| \leq 1$ ,

$j \in Z$ ,  $\mu_j = \text{sign}\{f_j(\mathbf{x}^*)\}$ ,  $j \notin Z$ . A quasi-Newton step is obtained by solving the linear system

$$\mathbf{J}_k \begin{bmatrix} \mathbf{h}_k \\ \Delta \delta_k \end{bmatrix} = -\mathbf{R}(\mathbf{x}_k, \delta_k), \quad (2.28)$$

where  $\mathbf{J}_k$  is an approximate Jacobian of (2.27) which consists of a mixture of the first derivatives  $\mathbf{f}'_z$  and approximations to the Hessians

$$\sum_{j \in Z} \mu_j \mathbf{f}'_j(\mathbf{x}) + \sum_{j \notin Z} \delta_j \mathbf{f}''_j(\mathbf{x}) \quad (2.29)$$

(Bandler, Kellermann and Madsen 1987). It is important to notice that in order to define the correct set of optimality equations, we must first identify the set of zero functions at the optimum. In practice, we typically use the current set of zero functions to construct and solve the optimality equations, known as the active set method. For this method to succeed, we must be sufficiently close to the solution. A similar concept applies to the minimax case in which the active set is defined by the worst error functions (Hald and Madsen 1981).

This example reveals that the application of quasi-Newton method can be quite involved when the ordinary gradient  $\mathbf{U}'$  does not exist. The theory of generalized gradient (Clarke 1975) addresses the optimality for a broad range of non-differentiable problems.

#### 2.4.2 Updates of a Hessian

Quasi-Newton methods, whether in (2.24) or (2.26), all require update of certain approximate Hessians. Among the formulas that have been proposed over the years, the most well-known are the

Powell Symmetric Broyden (PSB) update (Powell 1970b), the Davidon-Fletcher-Powell (DFP) update (Davidon 1959, Fletcher and Powell 1963) and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) update (Broyden 1969, Fletcher 1970, Goldfarb 1970, Shanno 1970). They are given by, respectively,

$$\begin{aligned} \mathbf{B}_{k+1}^{\text{PSB}} &= \mathbf{B}_k + \frac{\mathbf{w}\mathbf{s}^\top + \mathbf{s}\mathbf{w}^\top}{\mathbf{s}^\top \mathbf{s}} - \frac{\mathbf{w}^\top \mathbf{s} \mathbf{s}\mathbf{s}^\top}{(\mathbf{s}^\top \mathbf{s})^2}, \\ \mathbf{B}_{k+1}^{\text{DFP}} &= \mathbf{B}_k + \frac{\mathbf{w}\mathbf{y}^\top + \mathbf{y}\mathbf{w}^\top}{\mathbf{y}^\top \mathbf{s}} - \frac{\mathbf{w}^\top \mathbf{s} \mathbf{y}\mathbf{y}^\top}{(\mathbf{y}^\top \mathbf{s})^2}, \\ \mathbf{B}_{k+1}^{\text{BFGS}} &= \mathbf{B}_k + \frac{\mathbf{y}\mathbf{y}^\top}{\mathbf{y}^\top \mathbf{s}} - \frac{\mathbf{B}_k \mathbf{s}\mathbf{s}^\top \mathbf{B}_k}{\mathbf{s}^\top \mathbf{B}_k \mathbf{s}}, \end{aligned} \quad (2.30)$$

where  $\mathbf{s} = \mathbf{x}_{k+1} - \mathbf{x}_k$ ,  $\mathbf{y} = \mathbf{U}'(\mathbf{x}_{k+1}) - \mathbf{U}'(\mathbf{x}_k)$  (if  $\mathbf{B}$  is to approximate the Hessian of  $U$ ) or  $\mathbf{y} = \mathbf{f}'_j(\mathbf{x}_{k+1}) - \mathbf{f}'_j(\mathbf{x}_k)$  (if  $\mathbf{B}$  is to approximate  $\mathbf{f}'_j$ ), and  $\mathbf{w} = \mathbf{y} - \mathbf{B}_k \mathbf{s}$ . A thorough treatment of the theory underlying these updates has been given by Dennis and Moré (1977). As they have pointed out, numerical evidence seems to support the BFGS update as the best formula for use in minimization. The interesting expression

$$\mathbf{B}_{k+1}^\theta = \theta \mathbf{B}_{k+1}^{\text{DFP}} + (1 - \theta) \mathbf{B}_{k+1}^{\text{BFGS}} \quad (2.31)$$

describes the Broyden family (Broyden 1967, Fletcher 1970). Dixon (1972) has proved theoretically that when an exact line search is used all members of the Broyden family would have the same performance. In practice, the merits of a great many variations are often compared in terms of their preservation of positive definiteness, convergence to the true Hessian and numerical performance.

Sometimes it is more efficient to update the inverse of an

approximate Hessian using formulas similar to those in (2.30).

## 2.5 COMBINED METHODS

The Gauss-Newton and quasi-Newton methods each has its own advantage. The trust region Gauss-Newton method is globally convergent. But, like other first-order methods, it suffers from a very slow rate of convergence when close to a singular solution (Madsen 1985). On the other hand, the quasi-Newton method enjoys a fast rate of convergence near a solution but is not always reliable from a bad starting point. These complementary properties of the two methods seem to suggest their combination in one algorithm.

Hald and Madsen (1981, 1985) have developed a class of two-stage algorithms. A trust region Gauss-Newton method is employed in Stage 1 to provide global convergence to a neighbourhood of a solution. When the solution is singular, the first method suffers from a very slow rate of convergence and a switch is made to a quasi-Newton method (Stage 2). Several switches between the two methods may take place and the switching criteria ensure the global convergence of the combined algorithm. This approach has demonstrated a very strong performance in circuit optimization (Bandler, Kellermann and Madsen 1985, 1987).

Powell (1970c) has extended the Levenberg-Marquardt method and suggested a trust-region strategy which interpolates between a steepest descent step and a Newton step. When far away from the solution, the step is biased toward the steepest descent direction to make sure that it is downhill. Once near the solution, taking a

full Newton step will provide rapid final convergence.

## 2.6 AN ALGORITHM FOR LINEARLY CONSTRAINED ONE-SIDED $\ell_1$ OPTIMIZATION

### 2.6.1 Introductory Remarks

The linearly constrained one-sided  $\ell_1$  optimization problem to be considered has the following formulation.

$$\begin{aligned} & \underset{\mathbf{x}}{\text{minimize}} \quad U(\mathbf{x}) = \sum_{j \in J} f_j(\mathbf{x}) \\ & \text{subject to} \\ & \mathbf{a}_i^T \mathbf{x} + b_i = 0, \quad i = 1, 2, \dots, L_{e_q}, \\ & \mathbf{a}_i^T \mathbf{x} + b_i \geq 0, \quad i = L_{e_q} + 1, \dots, L, \end{aligned} \tag{2.32}$$

where  $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T$ ,  $J = \{j \mid f_j > 0\}$  identifies the set of positive functions among  $f_1, f_2, \dots, f_m$ ,  $\mathbf{a}_i$  and  $b_i$  are constants.

The problem arises in a number of applications. It can be applied to circuit design where  $f$  represents error functions arising from upper and lower specifications. Bandler, Kellermann and Madsen (1987) have considered multiplexer design by the one-sided  $\ell_1$  optimization. In Chapter 5 of this thesis, the one-sided  $\ell_1$  optimization constitutes an integrated part of the generalized  $\ell_p$  centering algorithm for yield enhancement.

Traditionally, the one-sided  $\ell_1$  problem is treated by defining

$$f_j^+ = \begin{cases} f_j, & \text{if } f_j > 0, \\ 0, & \text{otherwise,} \end{cases} \tag{2.33}$$

and minimizing the  $\ell_1$  norm of  $\mathbf{f}^+$ . This simplistic approach has two shortcomings. Firstly, the discontinuity of  $f_j^+$  at  $f_j = 0$  can be

and should be taken into account explicitly in the linear programming subproblem of each Gauss-Newton iteration. Introducing the discontinuity externally may affect adversely the local linearization of the error functions, which is essential to the Gauss-Newton method. Furthermore, the  $\ell_1$  and the one-sided  $\ell_1$  problems have different optimality conditions. Using an  $\ell_1$  algorithm to solve a one-sided  $\ell_1$  problem may lead to a false solution which satisfies the wrong set of optimality equations.

We now describe a one-sided  $\ell_1$  algorithm which is based on the Hald and Madsen approach.

### 2.6.2 Stage 1: A Trust Region Gauss-Newton Method

This is a direct application of the trust region Gauss-Newton method (Section 2.3) to the one-sided  $\ell_1$  problem.

At the  $k$ th iteration, a feasible point  $\mathbf{x}_k$  and a local bound  $\Lambda_k$  are given. The following subproblem is defined:

$$\begin{aligned}
 & \text{minimize}_{\mathbf{h}, \mathbf{y}} \quad \sum_{j=1}^m y_j \\
 & \text{subject to} \\
 & y_j \geq f_j(\mathbf{x}_k) + \mathbf{f}'_j(\mathbf{x}_k)^T \mathbf{h}, \quad j = 1, 2, \dots, m, \\
 & y_j \geq 0, \quad j = 1, 2, \dots, m, \\
 & \Lambda_k \geq h_i, \quad \Lambda_k \geq -h_i, \quad i = 1, 2, \dots, n, \\
 & \mathbf{a}_i^T(\mathbf{x}_k + \mathbf{h}) + b_i = 0, \quad i = 1, 2, \dots, L_{\text{eq}}, \\
 & \mathbf{a}_i^T(\mathbf{x}_k + \mathbf{h}) + b_i \geq 0, \quad i = L_{\text{eq}}+1, \dots, L,
 \end{aligned} \tag{2.34}$$

which can be solved by a standard linear programming routine. In the light of the discussions in Section 2.3, we can relate (2.34)

to a semi-linearized subproblem  $H(\bar{\mathbf{f}})$ . For each  $f_j$ , we define a piece-wise linearized model as

$$\bar{f}_j(\mathbf{h}) = \begin{cases} f_j + (\mathbf{f}'_j)^T \mathbf{h}, & \text{if } f_j + (\mathbf{f}'_j)^T \mathbf{h} > 0, \\ 0, & \text{otherwise,} \end{cases} \quad (2.35)$$

which corresponds to two linear constraints in (2.34). The discontinuity at  $\bar{f}_j = 0$  is built into the local model and handled internally by activating the appropriate constraint (both constraints are active at  $\bar{f}_j = 0$ ). In comparison, the simplistic approach (using the two-sided  $\ell_1$  and  $\mathbf{f}'_j$ ) assumes either  $\bar{f}_j = f_j + (\mathbf{f}'_j)^T \mathbf{h}$  or  $\bar{f}_j = 0$  throughout one iteration, depending on which side of  $f_j = 0$  the iteration starts. Such an assumption becomes invalid if the point of discontinuity is crossed during the iteration.

Denote the solution of (2.34) by  $\mathbf{h}_k$ . If  $\mathbf{x}_k + \mathbf{h}_k$  reduces the nonlinear objective function, i.e., if  $U(\mathbf{x}_k + \mathbf{h}_k) < U(\mathbf{x}_k)$ , then it is taken as the next iterate. Otherwise  $\mathbf{x}_{k+1} = \mathbf{x}_k$ .

The local bound  $\Lambda_k$  is adjusted in every iteration based on the goodness of the linearized model, following the general rules described in Section 2.3. More precisely, if

$$U(\mathbf{x}_k) - U(\mathbf{x}_k + \mathbf{h}_k) \leq 0.25[U(\mathbf{x}_k) - \bar{U}(\mathbf{h}_k)], \quad (2.36)$$

then the trust region appears to be too large and the bound is decreased:  $\Lambda_{k+1} = 0.25\Lambda_k$ . Otherwise, if

$$U(\mathbf{x}_k) - U(\mathbf{x}_k + \mathbf{h}_k) \geq 0.75[U(\mathbf{x}_k) - \bar{U}(\mathbf{h}_k)], \quad (2.37)$$

then the bound is increased:  $\Lambda_{k+1} = 2\Lambda_k$ . If neither (2.36) nor (2.37) holds then  $\Lambda_{k+1} = \Lambda_k$ .

### 2.6.3 Stage 2: A Quasi-Newton Method

The Stage 2 of the algorithm applies a quasi-Newton method to solving the optimality equations of the one-sided  $\ell_1$  problem.

The nonlinear programming equivalent to the linearly constrained one-sided  $\ell_1$  problem is

$$\begin{aligned} & \underset{\mathbf{x}, \mathbf{y}}{\text{minimize}} \quad U(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^m y_j \\ & \text{subject to} \\ & g_j = y_j - f_j(\mathbf{x}) \geq 0, \quad j = 1, 2, \dots, m, \\ & g_{j+m} = y_j \geq 0, \quad j = 1, 2, \dots, m, \\ & g_{j+2m} = \mathbf{a}_j^T \mathbf{x} + b_j \geq 0, \quad j = 1, 2, \dots, L. \end{aligned} \tag{2.38}$$

A local optimum must satisfy the Kuhn-Tucker conditions as

$$\frac{\partial U}{\partial(\mathbf{x}, \mathbf{y})} - \sum_{j=1}^{2m+L} \delta_j \frac{\partial g_j}{\partial(\mathbf{x}, \mathbf{y})} = \mathbf{0}, \tag{2.39}$$

where  $\delta_j$  are nonnegative multipliers.  $\delta_j = 0$  if  $g_j > 0$ .

First we examine the derivatives with respect to  $\mathbf{y}$ . It is obvious that

$$\begin{aligned} \partial U / \partial \mathbf{y} &= [1 \ 1 \ \dots \ 1]^T, \\ \partial g_j / \partial \mathbf{y} &= \mathbf{u}_j, \quad j = 1, 2, \dots, m, \\ \partial g_{j+m} / \partial \mathbf{y} &= \mathbf{u}_j, \quad j = 1, 2, \dots, m, \\ \partial g_{j+2m} / \partial \mathbf{y} &= \mathbf{0}, \quad j = 1, 2, \dots, L, \end{aligned} \tag{2.40}$$

where  $\mathbf{u}_j$  is the  $j$ th column vector of an  $M$  by  $M$  identity matrix. It has 1 in the  $j$ th position and zeros elsewhere. From (2.39) and (2.40) we have

$$[1 \ 1 \ \dots \ 1]^T - \sum_{j=1}^m (\delta_j \mathbf{u}_j + \delta_{j+m} \mathbf{u}_j) = \mathbf{0}. \tag{2.41}$$

The  $j$ th equation is actually

$$\delta_j + \delta_{j+m} = 1. \quad (2.42)$$

Notice that at least one of the two constraints corresponding to  $f_j$ , namely  $g_j$  and  $g_{j+m}$ , must be active. If  $f_j > 0$  then  $g_j$  is active but  $g_{j+m}$  is not, thus  $\delta_{j+m} = 0$  and  $\delta_j = 1$ . If, on the other hand,  $f_j < 0$ , then  $\delta_j = 0$ . When  $f_j = 0$  both  $g_j$  and  $g_{j+m}$  are active. We summarize these results as

$$\begin{aligned} \delta_j &= 1, & \text{if } j \in J, \\ 1 \geq \delta_j \geq 0, & & \text{if } j \in Z, \\ \delta_j &= 0, & \text{otherwise,} \end{aligned} \quad (2.43)$$

where  $J = \{j \mid f_j > 0\}$  and  $Z = \{j \mid f_j = 0\}$  are mutually exclusive index sets.

Now we inspect the derivatives with respect to  $\mathbf{x}$ . We have

$$\begin{aligned} \partial U / \partial \mathbf{x} &= \mathbf{0}, \\ \partial g_j / \partial \mathbf{x} &= -\mathbf{f}'_j(\mathbf{x}), & j = 1, 2, \dots, m, \\ \partial g_{j+m} / \partial \mathbf{x} &= \mathbf{0}, & j = 1, 2, \dots, m, \\ \partial g_{j+2m} / \partial \mathbf{x} &= \mathbf{a}_j, & j = 1, 2, \dots, L. \end{aligned} \quad (2.44)$$

Combining (2.39), (2.43) and (2.44), we arrive at the set of optimality equations as

$$\begin{aligned} \sum_{j \in J} \mathbf{f}'_j(\mathbf{x}) + \sum_{j \in Z} \delta_j \mathbf{f}'_j(\mathbf{x}) - \sum_{i \in A} \mu_i \mathbf{a}_i &= \mathbf{0}, \\ f_j(\mathbf{x}) &= 0, & j \in Z, \\ \mathbf{a}_i^T \mathbf{x} + b_i &= 0, & i \in A, \end{aligned} \quad (2.45)$$

where  $A = \{i \mid \mathbf{a}_i^T \mathbf{x} + b_i = 0\}$  identifies the active set of linear constraints which will always include the equality constraints,  $1 \geq \delta_j \geq 0$ ,  $j \in Z$ , and  $\mu_i = \delta_{i+2m} \geq 0$ . A local optimum of the linearly constrained one-sided  $\ell_1$  problem must satisfy (2.45).

Using a matrix notation, we express (2.45) as  $\mathbf{R}(\mathbf{x}, \boldsymbol{\delta}, \boldsymbol{\mu}) = \mathbf{0}$ .

The Stage 2 of our algorithm obtains a quasi-Newton step by solving the linear system

$$\mathbf{J}_k \begin{bmatrix} \Delta \mathbf{x}_k \\ \Delta \boldsymbol{\delta}_k \\ \Delta \boldsymbol{\mu}_k \end{bmatrix} = -\mathbf{R}(\mathbf{x}_k, \boldsymbol{\delta}_k, \boldsymbol{\mu}_k), \quad (2.46)$$

where  $\mathbf{J}_k$  is an approximation to the Jacobian of  $\mathbf{R}$  given by

$$\mathbf{R}' = \begin{bmatrix} \sum_{j \in J} \mathbf{f}'_j(\mathbf{x}) + \sum_{j \in Z} \delta_j \mathbf{f}'_j(\mathbf{x}) & \mathbf{E} & \mathbf{F} \\ & \mathbf{E}^T & \mathbf{0} \\ & -\mathbf{F}^T & \mathbf{0} \end{bmatrix} \quad (2.47)$$

where  $\mathbf{E}$  and  $\mathbf{F}$  are matrices with columns  $\mathbf{f}'_j(\mathbf{x})$ ,  $j \in Z$ , and  $-\mathbf{a}_i$ ,  $i \in A$ , respectively, and they can be calculated exactly at each iteration. The submatrix at the upper left-hand corner contains second-order derivatives, therefore, the corresponding part of  $\mathbf{J}_k$  is updated at each iteration using a modified BFGS formula (Powell 1978, also Bandler, Kellermann and Madsen 1987), as

$$\mathbf{B}_{k+1} = \mathbf{B}_k + \frac{\mathbf{y}\mathbf{y}^T}{\mathbf{y}^T \mathbf{s}} - \frac{\mathbf{B}_k \mathbf{s} \mathbf{s}^T \mathbf{B}_k}{\mathbf{s}^T \mathbf{B}_k \mathbf{s}}, \quad (2.48)$$

with

$$\begin{aligned} \mathbf{s} &= \Delta \mathbf{x}_k, \\ \mathbf{y} &= \theta [G(\mathbf{x}_k + \Delta \mathbf{x}_k, \boldsymbol{\delta}_k) - G(\mathbf{x}_k, \boldsymbol{\delta}_k)] + (1-\theta) \mathbf{B}_k \mathbf{s}, \\ G(\mathbf{x}, \boldsymbol{\delta}) &= \sum_{j \in J} \mathbf{f}'_j(\mathbf{x}) + \sum_{j \in Z} \delta_j \mathbf{f}'_j(\mathbf{x}), \end{aligned} \quad (2.48)$$

where  $\theta$  satisfying  $0 \leq \theta \leq 1$  is selected such that  $\mathbf{y}^T \mathbf{s} > 0$ , in order to maintain a positive definite  $\mathbf{B}_{k+1}$ .

#### 2.6.4 A Combined 2-Stage Algorithm

In the combined algorithm, the trust region Gauss-Newton method in Stage 1 is intended to provide global convergence and the quasi-Newton iteration of Stage 2 is used to obtain fast final convergence near a solution.

In Stage 1, the following steps take place.

Step 1 Given  $\mathbf{x}_k$  and  $\Lambda_k$ ,  $\mathbf{x}_{k+1}$  and  $\Lambda_{k+1}$  are found using the trust region Gauss-Newton method described in 2.6.2. The active sets are estimated as  $Z_{k+1}$  and  $A_{k+1}$  by the zero functions and active linear constraints at  $\mathbf{x}_{k+1}$ , respectively.

Step 2 An estimate  $(\delta_{k+1}, \mu_{k+1})$  of the multipliers is found through a least squares solution of (2.45) using  $\mathbf{x}_{k+1}$ ,  $Z_{k+1}$  and  $A_{k+1}$ . The approximate Jacobian is updated by the BFGS formula, giving  $J_{k+1}$ .

A switch from Stage 1 to Stage 2 is made if the following conditions are met.

- (a) The estimated active sets  $Z_{k+1}$  and  $A_{k+1}$  have been constant over  $K$  consecutive Stage 1 iterations (we use  $K = 3$ ).
- (b) The estimated multipliers corresponding to  $Z_{k+1}$  and  $A_{k+1}$  are in the correct ranges:  $1 \geq \delta_j \geq 0$  and  $\mu_i \geq 0$ .

The requirement of stable active sets and acceptable multipliers is intended to avoid premature switches to Stage 2.

The steps in Stage 2 have been described in the last section (2.6.3). A switch from Stage 2 back to Stage 1 is made if one of the following conditions holds.

- (a) The active sets are not complete because a function  $f_j$  with

$j \notin Z$  has become zero or changed sign, or a constraint not included in  $A$  has become violated.

- (b) The value of a multiplier is outside its range.
- (c) A quasi-Newton step fails to decrease the residual of the optimality equations:

$$\|R_{k+1}\| > 0.999\|R_k\|.$$

The use of the trust region Gauss-Newton method as Stage 1, the quasi-Newton method as Stage 2, and these switching conditions are at the heart of the class of algorithms described by Madsen (1985) who has proved global convergence in general.

The description of our algorithm for linearly constrained one-sided  $\ell_1$  optimization is complete. In Chapter 5, this algorithm is applied to circuit centering.

## 2.7 A UNIFIED FORTRAN LIBRARY FOR NONLINEAR OPTIMIZATION

Several optimization routines based on the Hald and Madsen approach have been implemented in a unified Fortran library which is named the KMOS library (Bandler, Chen and Renault 1987). It includes routines for linearly constrained minimax, linearly constrained  $\ell_1$ , linearly constrained one-sided  $\ell_1$  and unconstrained  $\ell_2$  optimization problems.

The routines for linearly constrained minimax problems were originally developed by Bandler and Zuberek (1982) for the CDC 170/730 system. Since then, many changes have taken place, both in the hardware and the software. Most importantly, many new algorithms have become available, including the  $\ell_1$  (Bandler, Kellermann

and Madsen 1987), the  $\ell_2$  (Madsen 1986) and the one-sided  $\ell_1$  (described in this Chapter) algorithms. Methods which do not require exact gradients have also been developed and implemented (which will be described in Chapter 7).

The KMOS library is created basically for the convenience of its use. The calling sequence to the optimization routines and the printing service provided by these routines are standardized and unified. To employ different optimization methods, only minimal changes need to be made to the user's program. By sharing some common codes, the unified library is also smaller in size as compared with the separate packages combined.

## 2.8 CONCLUDING REMARKS

In this chapter, we have addressed the recent advances in the state-of-the-art  $\ell_p$  optimization techniques. The formulation and the properties of the  $\ell_p$  functions have been reviewed. An important class of solution methods has been discussed in detail. We have shown that linear and linearly constrained linear minimax,  $\ell_2$  and  $\ell_1$  problems can be solved by linear or quadratic programming techniques. We have described the Gauss-Newton method which is based on the trust region solution of a semi-linearized subproblem and the quasi-Newton method which is applied to solving the set of optimality equations. The Levenberg-Marquardt method for the least squares problem has also been reviewed.

A new algorithm for the linearly constrained one-sided  $\ell_1$  optimization has been presented. The shortcomings of the tradi-

tional approach to the one-sided problems by externally defining discontinuous error functions have been exposed and, therefore, the need of a true one-sided algorithm was justified. We have defined the trust region Gauss-Newton iteration which solves a sequence of semi-linearized subproblems by linear programming. We have also derived the optimality equations for the one-sided  $\ell_1$  problem and applied a quasi-Newton method to the solution of these equations. Following the Hald and Madsen approach, the Gauss-Newton and the quasi-Newton methods have been combined into a 2-stage algorithm.

The optimization techniques which we have described provide the powerful tools for solving the various circuit optimization problems covered by the subsequent chapters. Due to their proven success in many practical applications, software based on these techniques has been integrated by EEs of Inc. into TOUCHSTONE Version 1.5 (1987).

## CHAPTER 3

### EFFICIENT APPROACHES TO NETWORK SENSITIVITY ANALYSIS

#### 3.1 INTRODUCTION

The application of gradient-based optimization techniques to circuit problems requires the evaluation of network sensitivities, typically first-order. Director and Rohrer (1969a, 1969b) pioneered the adjoint network approach to sensitivity analysis for linear circuits. Their work and subsequent contributions by many other researchers have greatly facilitated the advance in CAD tools from non-gradient (random, direct) techniques to sophisticated and powerful ones.

Relating to the gradient-based  $\ell_p$  optimization methods in Chapter 2, we are interested in the first-order derivatives of the functions  $f_1, f_2, \dots, f_m$  with respect to the variables  $x_1, x_2, \dots, x_n$ , i.e., the Jacobian of  $\mathbf{f}$  with respect to  $\mathbf{x}$ . For a circuit problem, the functions  $f_j$  are typically derived from the errors between the circuit responses and the given (constant) specifications. Therefore, the derivatives of  $f_j$  can be obtained from the appropriate network sensitivities.

In this chapter, we describe a unified and systematic approach to efficient sensitivity calculations for linear networks in the frequency domain. Useful formulas are derived from a nodal description of a linear network. Unterminated and terminated multi-ports are analyzed in general. The results are then applied

to two-ports which are widely used to represent filters as well as subnetworks. For two-ports, we also present some second-order sensitivities useful in the optimization of group delay and gain slope. Formulas for some commonly used frequency responses are also given.

A special class of two-ports, namely lossless two-ports, deserves separate attention because their sensitivity expressions can be shown in a simple analytic form. This result was stated by Orchard, Temes and Cataltepe for the first time in 1983. Three original and different proofs have been presented by Bandler, Chen and Daijavad (1984a, 1984b, 1985b). Here, we describe two of the proofs using a notation consistent with the rest of the chapter.

### 3.2 SENSITIVITY ANALYSIS USING A NODAL DESCRIPTION

For the network under consideration, assume that a nodal description is available. For simplicity, we further assume an admittance matrix. The formulas are, of course, applicable to an impedance or hybrid matrix.

We have

$$\mathbf{Y} \mathbf{V} = \mathbf{I} \quad (3.1)$$

where  $\mathbf{Y}$  is the  $N$  by  $N$  admittance matrix,  $\mathbf{V}$  the nodal voltages and  $\mathbf{I}$  the excitation vector. Differentiating (3.1) with respect to a generic variable  $\phi$  gives

$$\frac{\partial \mathbf{V}}{\partial \phi} = -\mathbf{Y}^{-1} \frac{\partial \mathbf{Y}}{\partial \phi} \mathbf{V}. \quad (3.2)$$

To select the sensitivity for a particular voltage of inte-

rest, say,  $V_k$ , we define a unit vector  $\mathbf{u}_k$  which is the  $k$ th column vector of an  $N$  by  $N$  identity matrix (its  $k$ th element is 1 and the others are zeros). Premultiply (3.2) by  $\mathbf{u}_k^T$ ,

$$\frac{\partial V_k}{\partial \phi} = -\mathbf{u}_k^T \mathbf{Y}^{-1} \mathbf{V} = -\hat{\mathbf{v}}^T \frac{\partial \mathbf{Y}}{\partial \phi} \mathbf{V} = -\sum_{i=1}^N \sum_{j=1}^N \frac{\partial Y_{ij}}{\partial \phi} \hat{v}_i V_j, \quad (3.3)$$

where we define an adjoint system by

$$\mathbf{Y}^T \hat{\mathbf{v}} = \mathbf{u}_k. \quad (3.4)$$

For example, consider a capacitor connected between nodes  $a$  and  $b$ . The parameter  $C$  appears in four places in  $\mathbf{Y}$ : as  $j\omega C$  in  $Y_{aa}$  and  $Y_{bb}$  and as  $-j\omega C$  in  $Y_{ab}$  and  $Y_{ba}$ . Therefore,

$$\frac{\partial V_k}{\partial C} = -j\omega(\hat{v}_a V_a + \hat{v}_b V_b - \hat{v}_a V_b - \hat{v}_b V_a) = -j\omega(\hat{v}_a - \hat{v}_b)(V_a - V_b). \quad (3.5)$$

If we solve the original circuit (3.1) by LU factorization, then the adjoint solution (3.4) requires minimal extra effort.

Director and Rohrer (1969a) have arrived at essentially the same results from Tellegen's theorem. The concise derivation by matrix algebra as shown here was due to Branin (1973).

### 3.3 GENERAL M-PORTS

#### 3.3.1 Unterminated M-Ports

It is quite common to separate the nodes of a network into internal and external nodes, and designate some pairs of external nodes to constitute external ports. This is especially useful in the analysis of a complicated system consisting of subnetworks. We can describe each subnetwork as an unterminated  $M$ -port and the rest of the system as terminations of the  $M$ -port. Such an equivalent of

the subnetwork may then be analyzed independently.

Suppose that the nodal equations (3.1) describe an unterminated network. We use

$$\begin{aligned} \mathbf{V}_P &= [V_{P1} \ V_{P2} \ \dots \ V_{PM}]^T, \\ \mathbf{I}_P &= [I_{P1} \ I_{P2} \ \dots \ I_{PM}]^T, \end{aligned} \quad (3.6)$$

to represent the voltages and currents, respectively, associated with the external ports, as shown in Fig. 3.1. Conventionally the unterminated M-port is characterized by

$$\mathbf{V}_P = \mathbf{z} \mathbf{I}_P, \quad (3.7)$$

where  $\mathbf{z}$  is the M-port open-circuit impedance matrix.

Consider a typical element of  $\mathbf{z}$ , say  $z_{k\ell}$ . Suppose that the  $k$ th port is created between nodes  $i$  and  $j$ , and  $V_{Pk} = V_i - V_j$ . By introducing a vector  $\mathbf{u}_{Pk}$  which has 1 in the  $i$ th position,  $-1$  in the  $j$ th position and zeros elsewhere, we can write

$$V_{Pk} = \mathbf{u}_{Pk}^T \mathbf{V}. \quad (3.8)$$

If node  $j$  is the ground then  $\mathbf{u}_{Pk}$  has 1 in the  $i$ th position and zeros elsewhere.  $\mathbf{u}_{P\ell}$  is similarly defined for the  $\ell$ th port.

It follows that

$$z_{k\ell} = \mathbf{u}_{Pk}^T \mathbf{Y}^{-1} \mathbf{u}_{P\ell} = \mathbf{u}_{Pk}^T \mathbf{p}_\ell, \quad (3.9)$$

where  $\mathbf{p}_\ell$  is the solution of

$$\mathbf{Y} \mathbf{p}_\ell = \mathbf{u}_{P\ell}. \quad (3.10)$$

The sensitivity formula for  $z_{k\ell}$  can be derived, similarly to the derivation of (3.3), to be

$$\frac{\partial z_{k\ell}}{\partial \phi} = -\hat{\mathbf{p}}_k^T \frac{\partial \mathbf{Y}}{\partial \phi} \mathbf{p}_\ell, \quad (3.11)$$

where  $\hat{\mathbf{p}}_k$  is the solution of

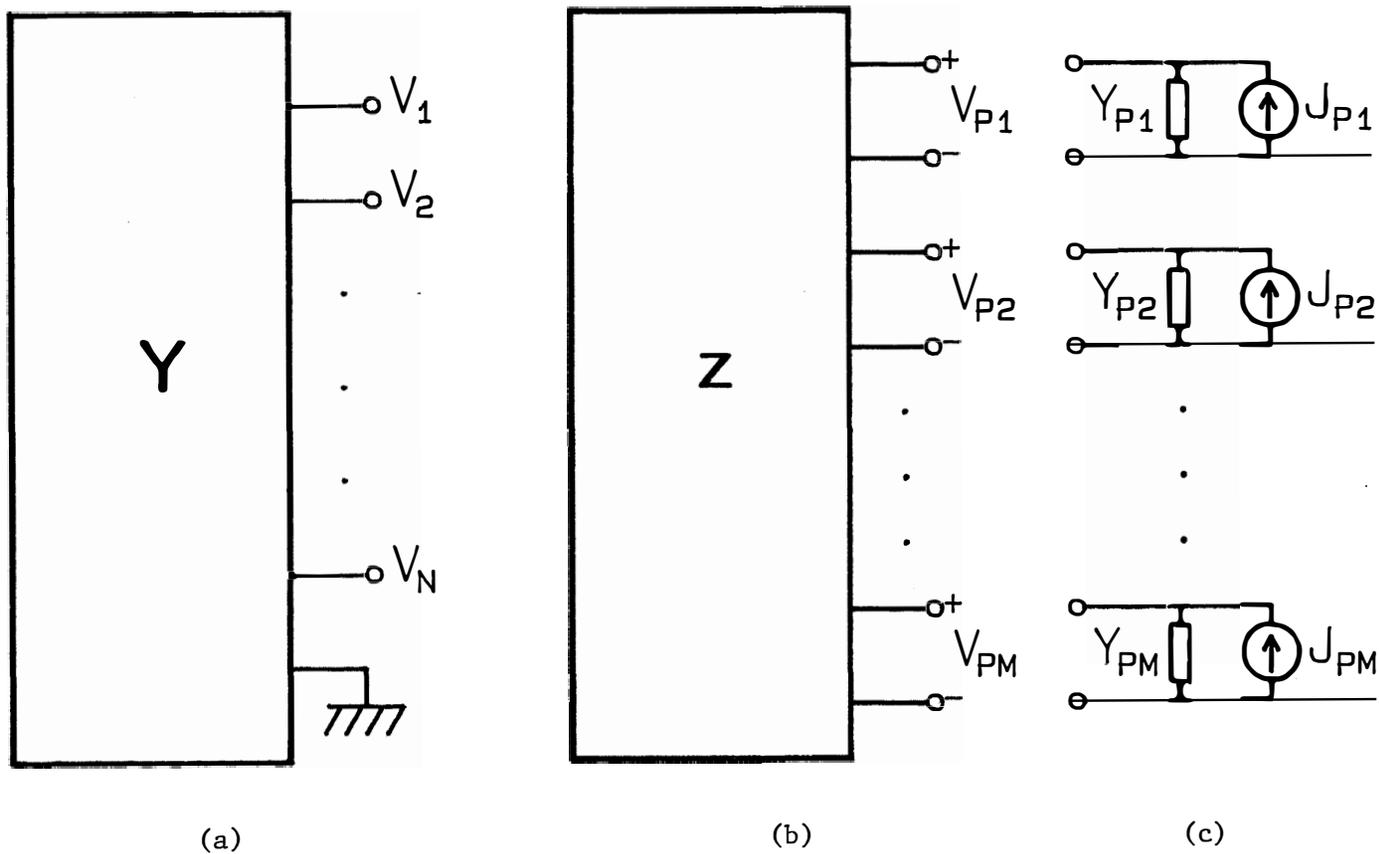


Fig. 3.1 An illustration of multi-ports. (a) An unterminated network described by a nodal admittance matrix. (b) The network represented by an unterminated M-port. (c) The terminations of the M-port.

$$\mathbf{Y}^T \hat{\mathbf{P}}_k = \mathbf{u}_{P_k}. \quad (3.12)$$

Generally, in order to complete the evaluation of  $z$  and its sensitivities, we need to solve the original network with  $M$  different right-hand sides, namely  $\mathbf{u}_{P_1}, \dots, \mathbf{u}_{P_M}$ , leading to  $\mathbf{P}_1, \dots, \mathbf{P}_M$ . We also need  $M$  adjoint solutions  $\hat{\mathbf{P}}_1, \dots, \hat{\mathbf{P}}_M$ . If the network is reciprocal then  $\mathbf{Y}^T = \mathbf{Y}$  and consequently  $\hat{\mathbf{P}}_k = \mathbf{P}_k$  for all  $k$ . While the use of an unterminated  $M$ -port allows us to analyze a network or subnetwork independent of its terminations, it also increases the amount of computations. We emphasize that all the solutions, original and adjoint, require only one LU factorization of  $\mathbf{Y}$ .

### 3.3.2 Terminated M-Ports

Now assume that the  $k$ th port is terminated by an independent current source  $J_{P_k}$  with an admittance  $Y_{P_k}$  (see Fig. 3.1). The currents and voltages of the  $M$ -ports are related by

$$\mathbf{I}_P = \mathbf{J}_P - \mathbf{Y}_P \mathbf{V}_P, \quad (3.13)$$

where

$$\mathbf{J}_P = [J_{P_1} \ J_{P_2} \ \dots \ J_{P_M}]^T, \quad (3.14)$$

$$\mathbf{Y}_P = \text{diag}(Y_{P_1} \ Y_{P_2} \ \dots \ Y_{P_M}).$$

Solving (3.7) and (3.13) concurrently gives

$$\mathbf{V}_P = (\mathbf{1} + z \mathbf{Y}_P)^{-1} z \mathbf{J}_P. \quad (3.15)$$

Differentiating (3.15) and after some algebra, we can arrive at a sensitivity formula for the terminated  $M$ -port as

$$\frac{\partial \mathbf{V}_P}{\partial \phi} = (\mathbf{1} + z \mathbf{Y}_P)^{-1} \left( \frac{\partial z}{\partial \phi} \mathbf{I}_P - z \frac{\partial \mathbf{Y}_P}{\partial \phi} \mathbf{V}_P \right). \quad (3.16)$$

### 3.4 TWO-PORTS

#### 3.4.1 General Two-Ports

A special case of M-ports, namely two-ports, is widely used to describe networks for which we are primarily interested in a pair of input and output variables. Conventionally the input port is defined between node 1 and the ground, and the output between node N and the ground. Thus,  $V_{P1} = V_1$  and  $V_{P2} = V_N$ .

Following the general approach for M-ports, we define two solutions of the original network as

$$\begin{aligned} \mathbf{Y} \mathbf{p} &= \mathbf{u}_1, \\ \mathbf{Y} \mathbf{q} &= \mathbf{u}_N, \end{aligned} \quad (3.17)$$

as well as two adjoint solutions as

$$\begin{aligned} \mathbf{Y}^T \hat{\mathbf{p}} &= \mathbf{u}_1, \\ \mathbf{Y}^T \hat{\mathbf{q}} &= \mathbf{u}_N, \end{aligned}$$

where  $\mathbf{u}_1 = [1 \ 0 \ \dots \ 0]^T$  and  $\mathbf{u}_N = [0 \ 0 \ \dots \ 0 \ 1]^T$ .

The open-circuit impedance matrix for the unterminated two-port is given by

$$\mathbf{z} = \begin{bmatrix} \mathbf{u}_1^T \mathbf{Y}^{-1} \mathbf{u}_1 & \mathbf{u}_1^T \mathbf{Y}^{-1} \mathbf{u}_N \\ \mathbf{u}_N^T \mathbf{Y}^{-1} \mathbf{u}_1 & \mathbf{u}_N^T \mathbf{Y}^{-1} \mathbf{u}_N \end{bmatrix} = \begin{bmatrix} P_1 & q_1 \\ P_N & q_N \end{bmatrix}, \quad (3.19)$$

and its sensitivities by

$$\frac{\partial \mathbf{z}}{\partial \phi} = -[\hat{\mathbf{p}} \ \hat{\mathbf{q}}]^T \frac{\partial \mathbf{Y}}{\partial \phi} [\mathbf{p} \ \mathbf{q}] = - \sum_{i=1}^N \sum_{j=1}^N \frac{\partial Y_{ij}}{\partial \phi} \begin{bmatrix} \hat{P}_i P_j & \hat{P}_i q_j \\ \hat{Q}_i P_j & \hat{Q}_i q_j \end{bmatrix}. \quad (3.20)$$

Usually, the output port is terminated by a load  $Y_L$ , and the input port by a source  $J = 1A$  with an admittance  $Y_S$ . These terminations are represented by

$$\begin{aligned} \mathbf{J}_P &= [1 \ 0]^T = \mathbf{u}_1, \\ \mathbf{Y}_P &= \begin{bmatrix} Y_S & 0 \\ 0 & Y_L \end{bmatrix}. \end{aligned} \quad (3.21)$$

Formulas (3.15) and (3.16) can then be applied directly to solving the terminated two-port and evaluating its sensitivities.

### 3.4.2 Second-Order Sensitivities

In the optimization of group delay and gain slope responses, second-order derivatives of  $\mathbf{z}$  are needed. In these cases, the first-order derivatives of  $\mathbf{z}$  with respect to the frequency  $\omega$  are used to evaluate the response itself. The sensitivities of  $\partial \mathbf{z} / \partial \omega$  with respect to a circuit parameter  $\phi$ , namely  $\partial^2 \mathbf{z} / \partial \omega \partial \phi$ , are of second-order.

Using (3.20) we have

$$\frac{\partial \mathbf{z}}{\partial \omega} = -[\hat{\mathbf{p}} \ \hat{\mathbf{q}}]^T \frac{\partial \mathbf{Y}}{\partial \omega} [\mathbf{p} \ \mathbf{q}] = - \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_N^T \end{bmatrix} \mathbf{Y}^{-1} \frac{\partial \mathbf{Y}}{\partial \omega} \mathbf{Y}^{-1} [\mathbf{u}_1 \ \mathbf{u}_N]. \quad (3.22)$$

Differentiating (3.22) with respect to  $\phi$  leads to

$$\begin{aligned} \frac{\partial^2 \mathbf{z}}{\partial \omega \partial \phi} &= \begin{bmatrix} \mathbf{u}^T \\ \mathbf{u}^T \end{bmatrix} \mathbf{Y}^{-1} \left( \frac{\partial \mathbf{Y}}{\partial \phi} \mathbf{Y}^{-1} \frac{\partial \mathbf{Y}}{\partial \omega} - \frac{\partial^2 \mathbf{Y}}{\partial \omega \partial \phi} + \frac{\partial \mathbf{Y}}{\partial \omega} \mathbf{Y}^{-1} \frac{\partial \mathbf{Y}}{\partial \phi} \right) \mathbf{Y}^{-1} [\mathbf{u}_1 \ \mathbf{u}_N] \\ &= [\bar{\mathbf{p}} \ \bar{\mathbf{q}}]^T \frac{\partial \mathbf{Y}}{\partial \phi} [\bar{\mathbf{p}} \ \bar{\mathbf{q}}] - [\hat{\mathbf{p}} \ \hat{\mathbf{q}}]^T \frac{\partial^2 \mathbf{Y}}{\partial \omega \partial \phi} [\mathbf{p} \ \mathbf{q}] + [\tilde{\mathbf{p}} \ \tilde{\mathbf{q}}]^T \frac{\partial \mathbf{Y}}{\partial \phi} [\mathbf{p} \ \mathbf{q}], \end{aligned} \quad (3.23)$$

where  $\bar{\mathbf{p}}$ ,  $\bar{\mathbf{q}}$ ,  $\tilde{\mathbf{p}}$ ,  $\tilde{\mathbf{q}}$  are solutions of, respectively,

$$\begin{aligned} \mathbf{Y} \bar{\mathbf{p}} &= [\partial \mathbf{Y} / \partial \omega] \mathbf{p}, \\ \mathbf{Y} \bar{\mathbf{q}} &= [\partial \mathbf{Y} / \partial \omega] \mathbf{q}, \\ \mathbf{Y}^T \tilde{\mathbf{p}} &= [\partial \mathbf{Y} / \partial \omega]^T \hat{\mathbf{p}}, \\ \mathbf{Y}^T \tilde{\mathbf{q}} &= [\partial \mathbf{Y} / \partial \omega]^T \hat{\mathbf{q}}. \end{aligned} \quad (3.24)$$

Eight solutions of the network are involved in (3.23) which require one LU factorization of  $\mathbf{Y}$  and eight forward and backward substitutions. Notice that four solutions will suffice for a reciprocal network.

### 3.4.3 Sensitivities of Commonly Used Frequency Responses

It is very common, especially in filter design, to use frequency responses, such as reflection coefficient, return loss, insertion loss, scattering parameters and group delay, to describe the external behaviour of a circuit. Once we have the sensitivities of the port voltages and currents, the corresponding formulas for the frequency responses can be readily derived. Table 3.1 summarizes the results for some commonly used frequency responses according to their conventional definitions.

### 3.4.4 Lossless Two-Ports

Lossless two-ports are widely used as prototype models in filter design. In the context of this chapter, lossless two-ports deserve a special treatment because we can prove that the related sensitivity expressions require only one network analysis.

The central sensitivity formulas for lossless two-ports were due to Orchard, Temes and Cataltepe (1983, 1985). Between 1983 when the formulas appeared for the first time and 1985 when their proof was presented by Orchard et al., three different and original proofs were published by Bandler, Chen and Daijavad (1984a, 1984b, 1985). Here, we present two proofs based on the

TABLE 3.1  
SENSITIVITY EXPRESSIONS FOR SELECTED FREQUENCY RESPONSES

Response	Formula	Sensitivity Expression
input reflection coefficient $\rho_{in}$	$\frac{Y_S}{Y_S^*} [2G_S V_1 - 1]$	$2[V_1 \frac{\partial}{\partial \phi} (\frac{Y_S G_S}{Y_S^*}) + \frac{Y_S G_S}{Y_S^*} \frac{\partial V_1}{\partial \phi}]$
input return loss	$-20 \log_{10}  \rho_{in} $	$-\frac{20}{\ln 10} \operatorname{Re} [\frac{1}{\rho_{in}} \frac{\partial \rho_{in}}{\partial \phi}]$
insertion loss	$-20 \log_{10}  V_N Y_T $	$-\frac{20}{\ln 10} \operatorname{Re} [\frac{1}{V_N} \frac{\partial V_N}{\partial \phi} + \frac{1}{Y_T} \frac{\partial Y_T}{\partial \phi}]$
group delay	$-\operatorname{Im} [\frac{1}{V_N} \frac{\partial V_N}{\partial \omega} + \frac{1}{Y_S} \frac{\partial Y_S}{\partial \omega}]$	$-\operatorname{Im} [\frac{1}{V_N} \frac{\partial^2 V_N}{\partial \omega \partial \phi} - \frac{1}{V_N^2} \frac{\partial V_N}{\partial \omega} \frac{\partial V_N}{\partial \phi} + \frac{1}{Y_S} \frac{\partial^2 Y_S}{\partial \omega \partial \phi} - \frac{1}{Y_S^2} \frac{\partial Y_S}{\partial \omega} \frac{\partial Y_S}{\partial \phi}]$
scattering matrix	$(\bar{z} - 1)(\bar{z} + 1)^{-1}$	$\frac{1}{2Z_0} (1 - S) \frac{\partial z}{\partial \phi} (1 - S)$

$G_S = \operatorname{Re}(Y_S)$ ,  $Y_T = Y_S + Y_L$ ,  $Y_S^*$  is the complex conjugate of  $Y_S$ .

$\bar{z} = z/Z_0$  where  $Z_0$  is the normalizing impedance.

$S$  is the scattering matrix.

ideas of the previous publications but following the notation consistently used in this chapter. The first proof is derived through algebraic manipulations. The second proof is based on the principle of conservation of energy and the Cauchy-Riemann equations of complex differentiation. It bears clear physical interpretation and mathematical elegance.

Theorem

Assuming that a lossless two-port is terminated by a source  $J = 1A$  with a conductance  $G_S$  at the input port and by a load  $G_L$  at the output port, the evaluation of the sensitivities of the transducer coefficient which is defined by

$$\theta = - \ln(2V_N \sqrt{G_S G_L}) \quad (3.25)$$

requires only one solution of the network.

Proof 1

Define

$$Y' = G_S u_1 u_1^T + G_L u_N u_N^T + Y \quad (3.26)$$

as the admittance matrix of the network including the terminations.

The reciprocal and lossless properties of the network implies

$$\begin{aligned} Y^T &= Y, \\ Y^* &= -Y. \end{aligned} \quad (3.27)$$

where \* stands for complex conjugate. The terminated network can be solved from

$$Y' \mathbf{v} = \mathbf{u}_1. \quad (3.28)$$

From (3.3) and noticing that  $Y'$  is symmetrical, we have

$$\frac{\partial V_1}{\partial \phi} = - \mathbf{v}^T \frac{\partial Y'}{\partial \phi} \mathbf{v}. \quad (3.29)$$

Differentiate (3.28) and premultiply the result by  $(\mathbf{V}^*)^T$ ,

$$(\mathbf{V}^*)^T \mathbf{Y}' \frac{\partial \mathbf{V}}{\partial \phi} = - (\mathbf{V}^*)^T \frac{\partial \mathbf{Y}'}{\partial \phi} \mathbf{V}. \quad (3.30)$$

Using (3.26) and (3.27) we have

$$\begin{aligned} \mathbf{Y}' + [(\mathbf{Y}')^*]^T &= 2(\mathbf{G}_S \mathbf{u}_1 \mathbf{u}_1^T + \mathbf{G}_L \mathbf{u}_N \mathbf{u}_N^T) + \mathbf{Y} + (\mathbf{Y}^*)^T \\ &= 2(\mathbf{G}_S \mathbf{u}_1 \mathbf{u}_1^T + \mathbf{G}_L \mathbf{u}_N \mathbf{u}_N^T). \end{aligned} \quad (3.31)$$

Evaluate  $\mathbf{Y}'$  from (3.31) and substitute the result for (3.30),

$$(\mathbf{V}^*)^T \{2(\mathbf{G}_S \mathbf{u}_1 \mathbf{u}_1^T + \mathbf{G}_L \mathbf{u}_N \mathbf{u}_N^T) - [(\mathbf{Y}')^*]^T\} \frac{\partial \mathbf{V}}{\partial \phi} = -(\mathbf{V}^*)^T \frac{\partial \mathbf{Y}'}{\partial \phi} \mathbf{V}. \quad (3.32)$$

From (3.28) we notice that  $(\mathbf{V}^*)^T [(\mathbf{Y}')^*]^T = \mathbf{u}_1^T$ , therefore

(3.32) can be reduced to

$$2\mathbf{G}_S \mathbf{V}_1^* \frac{\partial \mathbf{V}_1}{\partial \phi} + 2\mathbf{G}_L \mathbf{V}_N^* \frac{\partial \mathbf{V}_N}{\partial \phi} - \frac{\partial \mathbf{V}_1}{\partial \phi} = -(\mathbf{V}^*)^T \frac{\partial \mathbf{Y}'}{\partial \phi} \mathbf{V}. \quad (3.33)$$

By substituting (3.29) in (3.33) and using the input reflection coefficient  $\rho = 2\mathbf{G}_S \mathbf{V}_1 - 1$  (Table 3.1), we arrive at

$$\frac{\partial \mathbf{V}_N}{\partial \phi} = \frac{1}{2\mathbf{G}_L \mathbf{V}_N^*} (\rho^* \mathbf{V} - \mathbf{V}^*)^T \frac{\partial \mathbf{Y}'}{\partial \phi} \mathbf{V}. \quad (3.34)$$

Finally, from the definition of the transducer coefficient as given in (3.25), we obtain

$$\frac{\partial \theta}{\partial \phi} = - \frac{1}{\mathbf{V}_N} \frac{\partial \mathbf{V}_N}{\partial \phi} = \frac{1}{2\mathbf{P}_N} (\mathbf{V}^* - \rho^* \mathbf{V})^T \frac{\partial \mathbf{Y}'}{\partial \phi} \mathbf{V}, \quad (3.35)$$

where  $\mathbf{P}_N$  is the power in the load, given by  $\mathbf{P}_N = \mathbf{G}_L \mathbf{V}_N \mathbf{V}_N^*$ . Clearly, (3.35) involves only one solution, namely  $\mathbf{V}$ , of (3.28).

For example, let  $\phi$  be  $y_{ij}$  (the admittance connected between nodes  $i$  and  $j$ ). Then

$$\partial \mathbf{Y}' / \partial y_{ij} = (\mathbf{u}_i - \mathbf{u}_j)(\mathbf{u}_i - \mathbf{u}_j)^T, \quad (3.36)$$

and consequently,

$$\begin{aligned} \frac{\partial \theta}{\partial y_{ij}} &= \frac{1}{2P_N} [(V_i^* - \rho^* V_i) - (V_j^* - \rho^* V_j)] (V_i - V_j) \\ &= \frac{|V_{ij}|^2 - \rho^* V_{ij}^2}{2P_N}, \end{aligned} \quad (3.37)$$

where  $V_{ij} = V_i - V_j$  is the voltage across  $y_{ij}$ .

### Proof 2

Consider an internal branch between nodes  $i$  and  $j$  as characterized by  $I_{ij} = y_{ij} V_{ij}$ , where  $y_{ij} = g_{ij} + jx_{ij}$  and  $g_{ij} = 0$  at nominal (i.e.,  $y_{ij}$  represents a lossless element). The real power associated with this branch is given by  $P_{ij} = g_{ij} |V_{ij}|^2$ , which is equal to zero at nominal. Denote the power in the load by  $P_N$  and define

$$P_1 = \text{Re}[-V_1^* I_1] = \text{Re}[V_1^* (V_1 G_S - 1)]. \quad (3.38)$$

The conservation of energy of the whole system implies

$$P_1 + P_N + \sum P_{ij} = P_1 + P_N + \sum g_{ij} |V_{ij}|^2 = 0, \quad (3.39)$$

where the summation is taken over all the internal branches.

Differentiating (3.39) with respect to  $g_{ij}$  and  $x_{ij}$ , we have

$$\frac{\partial P_1}{\partial g_{ij}} + \frac{\partial P_N}{\partial g_{ij}} + |V_{ij}|^2 = 0, \quad (3.40)$$

and, at  $g_{ij} = 0$ ,

$$\frac{\partial P_1}{\partial x_{ij}} + \frac{\partial P_N}{\partial x_{ij}} = 0. \quad (3.41)$$

From (3.38), we obtain, after simple manipulations,

$$\frac{\partial P_1}{\partial \phi} = \text{Re}[(2G_S V_1^* - 1) \frac{\partial V_1}{\partial \phi}] = \text{Re}[\rho^* \frac{\partial V_1}{\partial \phi}]. \quad (3.42)$$

Following (3.29) and (3.36), we have  $\partial V_1 / \partial g_{ij} = -V_{ij}^2$  and  $\partial V_1 / \partial x_{ij} = -jV_{ij}^2$ . Therefore,

$$\frac{\partial P_1}{\partial g_{ij}} - j \frac{\partial P_1}{\partial x_{ij}} - \operatorname{Re}[\rho^* V_{ij}^2] + j \operatorname{Re}[j\rho^* V_{ij}^2] = -\rho^* V_{ij}^2. \quad (3.43)$$

Combine (3.40), (3.41) and (3.43),

$$\frac{\partial P_N}{\partial g_{ij}} - j \frac{\partial P_N}{\partial x_{ij}} - |V_{ij}|^2 + \rho^* V_{ij}^2. \quad (3.44)$$

The complex valued transducer coefficient given in (3.25) is analytical in the network parameters wherever it is defined. Let it be  $\theta = \alpha + j\beta$ . With respect to  $y_{ij} = g_{ij} + jx_{ij}$ , we know that the Cauchy-Riemann equations are satisfied as

$$\begin{aligned} \frac{\partial \alpha}{\partial g_{ij}} &= \frac{\partial \beta}{\partial x_{ij}}, \\ \frac{\partial \alpha}{\partial x_{ij}} &= -\frac{\partial \beta}{\partial g_{ij}}. \end{aligned} \quad (3.45)$$

Accordingly we find that (Lang 1977)

$$\frac{\partial \theta}{\partial y_{ij}} = \frac{\partial \alpha}{\partial g_{ij}} - j \frac{\partial \alpha}{\partial x_{ij}}. \quad (3.46)$$

By definition,

$$\begin{aligned} \alpha &= \operatorname{Re}[-\ln(2V_N \sqrt{G_S G_L})] = -(1/2) \ln(4|V_N|^2 G_S G_L) \\ &= -(1/2) [\ln(4G_S) + \ln(P_N)], \end{aligned} \quad (3.47)$$

therefore,  $\partial \alpha / \partial \phi = -[\partial P_N / \partial \phi] / 2P_N$ . It follows that

$$\frac{\partial \theta}{\partial y_{ij}} = -\frac{1}{2P_N} \left[ \frac{\partial P_N}{\partial g_{ij}} - j \frac{\partial P_N}{\partial x_{ij}} \right] = \frac{|V_{ij}|^2 - \rho^* V_{ij}^2}{2P_N}. \quad (3.48)$$

This formula is identical to (3.37).

### 3.5 CONCLUDING REMARKS

In this chapter, we have described a unified and systematic approach to efficient sensitivity calculations for linear network in the frequency domain. Useful formulas have been derived for networks described by a nodal matrix as well as unterminated and terminated multi-ports. A more elaborate treatment has been directed at two-ports including results for second-order sensitivity and frequency responses. Two elegant proofs have been derived for an important sensitivity expression related to lossless two-ports.

We recognize that exact sensitivity expressions are not always available, e.g., when time-domain analysis and nonlinear circuits are involved. The subject of gradient approximations is treated in Chapter 7.

Even for linear circuits in the frequency domain, large-scale networks present new problems which need to be addressed. Often, a large network can be described through compounded and interconnected subnetworks. Many commercial CAD programs such as SUPER-COMPACT (1986) and TOUCHSTONE (1985) have facilitated such a block structure.

One possible approach is to assemble the overall nodal matrix and solve the network equations using sparse techniques. Another approach is to rearrange the overall nodal matrix into a bordered block structure which is then solved using decomposition techniques (Hachtel and Sangiovanni-Vincentelli 1981). Also, it may be possible to develop efficient formulas for a special structure, such as the work of Bandler, Daijavad and Zhang (1986) for

multiplexing networks.

In conjunction with these techniques, the sensitivity analyses outlined in this chapter can be carried out at the subnetwork level.

## CHAPTER 4

### NOMINAL CIRCUIT OPTIMIZATION

#### 4.1 INTRODUCTION

Nominal circuit design is an approach to the optimization of a single point, in the space of designable variables, which best meets a given set of performance specifications. The classical paper by Temes and Calahan in 1967 was one of the earliest to formally advocate the use of iterative optimization in circuit design. Also, pioneering papers by Lasdon, Suchman and Waren (1966), Waren, Lasdon and Suchman (1967) demonstrated optimal design of linear arrays and filters using the penalty function approach. Since then optimization oriented CAD techniques have become indispensable tools in many engineering fields.

At the heart of the problem is the mathematical description of the engineering system under consideration and the design goals. In this chapter, we first introduce a hierarchy of models which are employed to simulate the physical system. The parameters and the response functions associated with these models are identified. Error functions which arise from the performance specifications and the simulated responses are defined. The objective function and variables for optimization are then discussed. The formulation of a nominal circuit optimization problem, at increasingly abstract levels, is clarified in our presentation.

Practical illustrations of nominal design are provided by

the optimization of multi-coupled cavity filters which are of significant interest in satellite communication systems. The large variety and complexity of these filters have made them prime candidates for computer-aided design. An efficient approach to the exact sensitivity analysis of multi-coupled cavity filters has been presented by Bandler, Chen and Daijavad (1986a). In this chapter, three examples of filter design are described, including a 10th-order elliptic filter, a 10th-order self-equalized filter obtained from simultaneous optimization of the amplitude and group delay, as well as a 6th-order asymmetric design. First-order prediction of the effect of cavity dissipation using filter sensitivities is also discussed.

A 16-channel multiplexer consisting of cavity filters distributed along a waveguide manifold which involves 240 nonlinear variables is also presented to illustrate efficient solutions to large-scale nominal optimization problems.

## 4.2 BASIC FORMULATION

### 4.2.1 A Hierarchy of Simulation Models

In order to apply the mathematical tools to an engineering problem, we have to be able to describe the physical system under consideration by suitable simulation models.

The physical system being described can be a network, a device, a process, and so on, which has a fixed structure and given element types. We manipulate the system through some adjustable parameters denoted by a column vector  $\phi^M$ . We use the superscript M

to identify concepts related to the physical system. Geometrical dimensions such as the width of a strip and the length of a waveguide section are examples of adjustable parameters.

In the production of integrated circuits,  $\phi^M$  may include some fundamental variables which control, for instance, a photomasking or doping process and, consequently, determine the geometrical and electrical parameters of a chip. External controls, such as the biasing voltages applied to an active device, may also be candidates for  $\phi^M$ .

The performance and characteristics of the system are usually described in terms of some measurable quantities. Frequency and transient responses are typical examples. These measured responses, or simply measurements, are denoted by  $F^M(\phi^M)$ .

Simulation models can be usefully defined at many levels. Tromp (1977, 1978) has considered an arbitrary number of levels (also see Bandler, El-Kady, Kellermann and Zuberek 1983). Such a hierarchical definition often clarifies the concepts of tolerance and model uncertainty (which will be treated in Chapter 5).

For simplicity, we consider a hierarchy of models consisting of four typical levels as

$$\begin{aligned} F^H &= F^H(F^L), \\ F^L &= F^L(\phi^H), \\ \phi^H &= \phi^H(\phi^J). \end{aligned} \tag{4.1}$$

$\phi^J$  is a set of low-level model parameters. It is supposed to represent, as closely as possible, the adjustable parameters in the actual system, i.e.,  $\phi^M$ .

$\phi^H$  defines a higher-level model, typically an equivalent circuit, with respect to a fixed topology. Usually, the reason for using an equivalent circuit is the convenience of its analysis. The relationship between  $\phi^L$  and  $\phi^H$  is either derived from theory or given by a set of empirical formulas.

Next on the hierarchy we define the model responses at two possible levels. The low-level external representation, denoted by  $F^L$ , can be the frequency-dependent scattering parameters, unterminated y-parameters, transfer function coefficients, and so on. Although these quantities may or may not be directly measurable, they are very often used to represent a subsystem.

The high-level responses  $F^H$  directly correspond to the measured responses, namely  $F^M$ , which may be, e.g., the frequency responses such as return loss, insertion loss and group delay of a suitably terminated circuit.

A realistic example of a one-section transformer on strip-line was originally considered by Bandler, Liu and Tromp (1976b). The circuits and parameters, physical as well as model, are shown in Fig. 4.1. The physical parameters  $\phi^M$  (and the low-level model  $\phi^L$ ) include strip widths, section lengths, dielectric constants, strip and substrate thicknesses. The equivalent circuit parameters, denoted by  $\phi^H$ , include the effective linewidths, junction parasitic inductances and effective section length. The scattering matrix of the circuit with respect to idealized (matched) terminations can be a candidate for a low-level external represen-

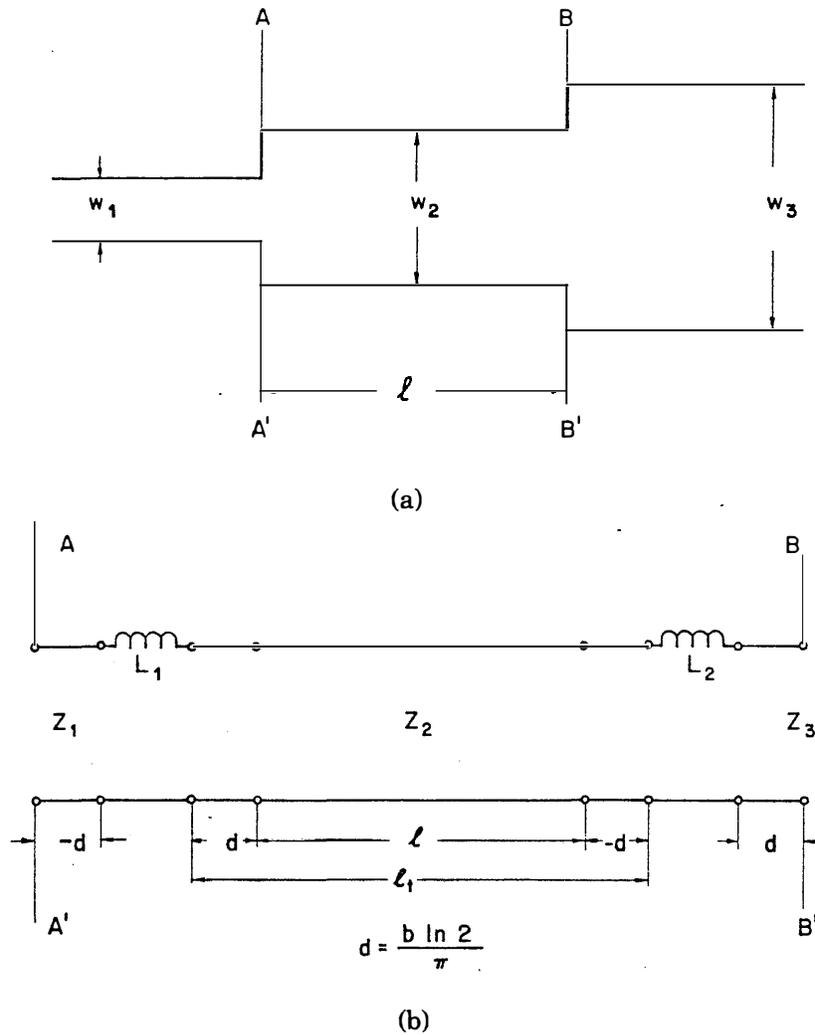


Fig. 4.1 A microwave stripline transformer (Bandler, Liu and Tromp 1976b) showing (a) the physical structure and (b) the equivalent circuit model. The physical parameters are

$$\phi^M = [w_1 \ w_2 \ w_3 \ \ell \ \sqrt{\epsilon_{r1}} \ \sqrt{\epsilon_{r2}} \ \sqrt{\epsilon_{r3}} \ b_1 \ b_2 \ b_3 \ t_{s1} \ t_{s2} \ t_{s3}]^T$$

where  $w$  is the strip width,  $\ell$  the length of the middle section,  $\epsilon_r$  the dielectric constant,  $b$  the substrate thickness and  $t_s$  the strip thickness.  $\phi^M$  is represented in the simulation model by  $\phi^L$ . The high-level parameters of the equivalent circuit are

$$\phi^H = [D_1 \ D_2 \ D_3 \ L_1 \ L_2 \ \ell_t]^T$$

where  $D$  is the effective linewidth,  $\ell_t$  the effective section length and  $L$  the junction parasitic inductance.

tation ( $F^L$ ). The reflection coefficient, taking into account the actual complex terminations, could be a high-level response of interest ( $F^H$ ).

For a particular case, we may choose a certain section of this hierarchy to form a design problem. We can choose either  $\phi^L$  or  $\phi^H$  as the designable parameters. Either  $F^L$  or  $F^H$  or a suitable combination of the both may be selected as the response functions. Bearing this in mind, we simplify the notation by using  $\phi$  for the designable parameters and  $F$  for the response functions.

#### 4.2.2 Specifications and Error Functions

We express the desirable performance of the system by a set of specifications which are usually functions of some independent variable(s) such as frequency, time, temperature, etc. (Bandler and Rizk 1979). In practice, we have to consider a discrete set of samples of the independent variable(s) such that satisfying the specifications at these points implies satisfying them almost everywhere. Also, we may consider simultaneously more than one kind of responses. Thus, without loss of generality, we denote a set of sampled specifications and the corresponding set of calculated response functions by, respectively,

$$\begin{aligned} S_j, & \quad j = 1, 2, \dots, m, \\ F_j(\phi), & \quad j = 1, 2, \dots, m. \end{aligned} \tag{4.2}$$

Error functions arise from the difference between the given specifications and the calculated responses. To formulate the error functions properly, we may wish to distinguish between having

upper and lower specifications (windows) and having single specifications, as illustrated, respectively, in Figs. 4.2(a) and 4.3(a). Sometimes the one-sidedness of upper and lower specifications is quite obvious, as in the design of a bandpass filter. On other occasions the distinction is more subtle, since a single specification may as well be interpreted as a window having zero width.

In the case of having single specifications, we define the error functions by

$$e_j(\boldsymbol{\phi}) = w_j |F_j(\boldsymbol{\phi}) - S_j|, \quad j = 1, 2, \dots, m, \quad (4.3)$$

where  $w_j$  is a nonnegative weighting factor.

In the case of having an upper specification  $S_{uj}$  and a lower specification  $S_{lj}$ , we define the error functions as

$$\begin{aligned} e_{uj}(\boldsymbol{\phi}) &= w_{uj} (F_j(\boldsymbol{\phi}) - S_{uj}), & j \in J_u, \\ e_{lj}(\boldsymbol{\phi}) &= w_{lj} (F_j(\boldsymbol{\phi}) - S_{lj}), & j \in J_l, \end{aligned} \quad (4.4)$$

where  $w_{uj}$  and  $w_{lj}$  are nonnegative weighting factors. The index sets as defined by

$$\begin{aligned} J_u &= \{j_1, j_2, \dots, j_k\}, \\ J_l &= \{j_{k+1}, j_{k+2}, \dots, j_m\} \end{aligned} \quad (4.5)$$

are not necessarily disjoint (i.e., we may have simultaneous upper and lower specifications). In order to have a set of uniformly indexed error functions, we let

$$\begin{aligned} e_i &= e_{uj}(\boldsymbol{\phi}), & j = j_i, & i = 1, 2, \dots, k, \\ e_i &= -e_{lj}(\boldsymbol{\phi}), & j = j_i, & i = k+1, k+2, \dots, m. \end{aligned} \quad (4.6)$$

The responses corresponding to the single specifications can be real or complex whereas upper and lower specifications can only be defined for real responses. Notice that in either case

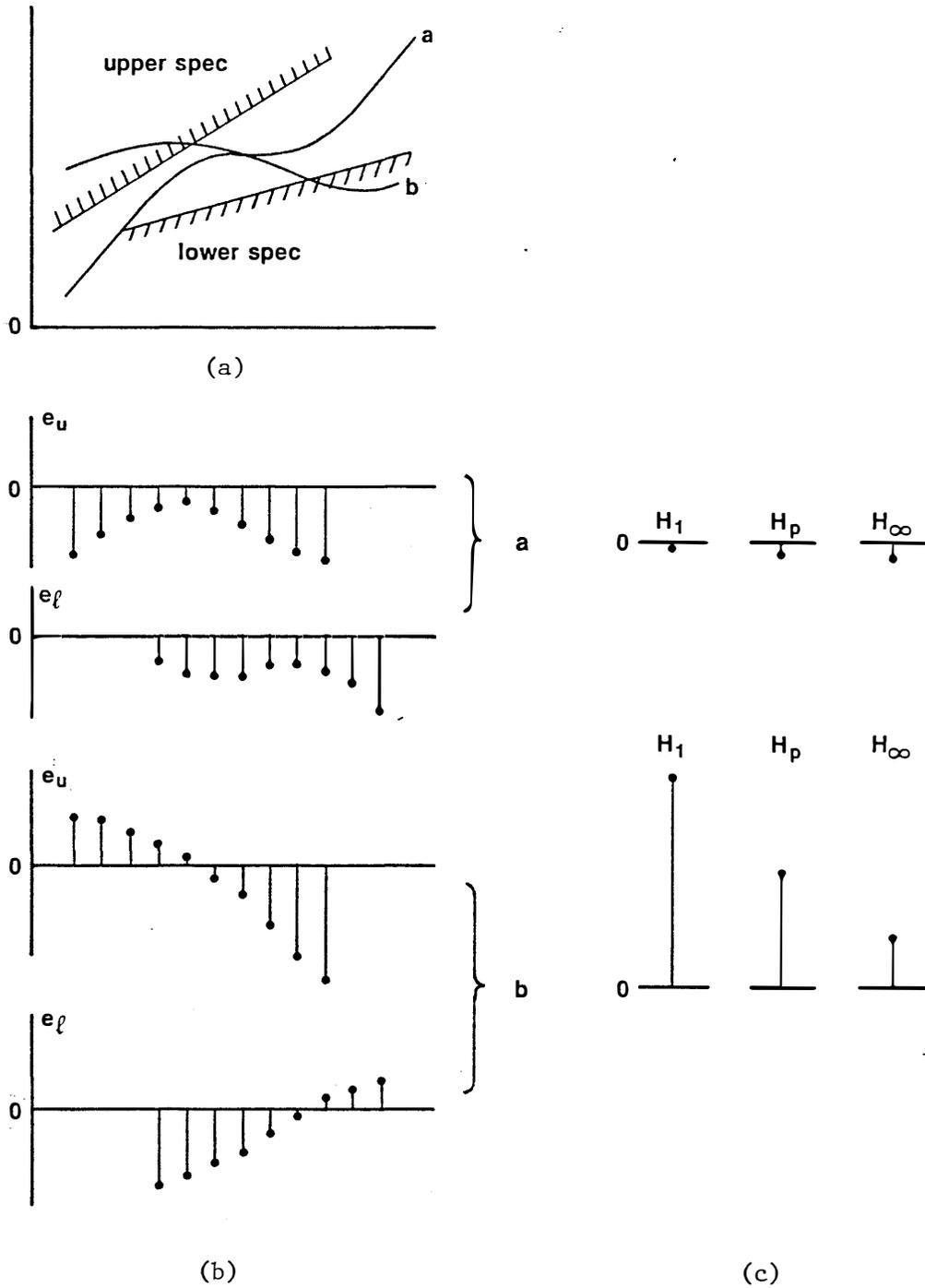


Fig. 4.2 Illustrations of (a) upper specifications, lower specifications and the responses of circuit a and circuit b, (b) error functions corresponding to circuits a and b, and (c) generalized  $\ell_p$  objective functions.

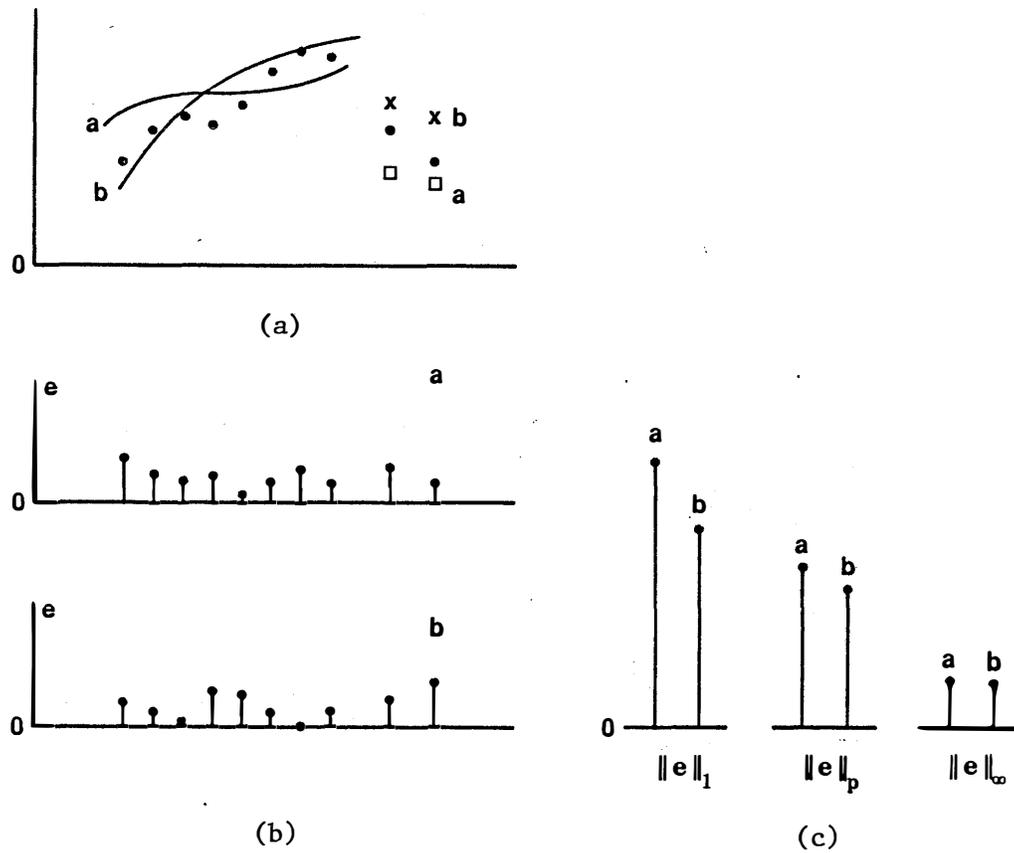


Fig. 4.3 Illustrations of (a) a discretized single specification and two discrete single specifications (e.g., expected parameter values to be matched), and the responses of circuit a and circuit b, (b) error functions related to circuits a and b, and (c) the corresponding  $l_p$  norms.

the error functions are real. Clearly, a positive (nonpositive) error function indicates a violation (satisfaction) of the corresponding specification. Figs. 4.2(b) and 4.3(b) depict the error functions corresponding to upper/lower specifications and single specifications, respectively.

#### 4.2.3 Variables and Objective Functions

In a nominal design, without considering tolerances (i.e., assuming that modeling and manufacturing can be done with absolute accuracy), we seek a single set of parameters, called a nominal point and denoted by  $\phi^0$ , which best satisfies the specifications. Furthermore, if the functional relationship of  $\phi^H = \phi^H(\phi^L)$  is considered to be precise, then it does not really matter at which level the design is conceived. In fact, traditionally it is often oriented to an equivalent circuit.

A classical case is network synthesis where  $\phi^{H,0}$  is obtained through the use of an equivalent circuit and/or a transfer function. A low-level model  $\phi^{L,0}$  is then calculated from  $\phi^{H,0}$ , typically with the help of an empirical formula (e.g., the number of turns of a coil is calculated for a given inductance). Finally, we try to realize  $\phi^{L,0}$  by its physical counterpart  $\phi^{M,0}$ .

With the tool of mathematical optimization, the nominal design  $\phi^0$  (at a chosen level) can be obtained through the solution of the following problem

$$\underset{\mathbf{x}}{\text{minimize}} U(\mathbf{x}), \quad (4.7)$$

where  $\mathbf{x}$  is a set of optimization variables and  $U(\mathbf{x})$  is a scalar

objective function which is typically defined as an  $\ell_p$  function  $H(\mathbf{e})$ , such as the ubiquitous least squares, the more esoteric generalized  $\ell_p$  or the minimax objective, as has been discussed in Chapter 2 and is depicted in Figs. 4.2(c) and 4.3(c).

Optimization variables and model parameters are in fact two separate concepts. The vector  $\mathbf{x}$  may contain all the elements or a subset of the elements of  $\boldsymbol{\phi}$ . It is a common practice to have some of the variables normalized. It is also common to have several model parameters tied to a single variable. Such dependencies usually exist in symmetrical circuit structures but, most importantly, they become a fact of life in integrated circuits.

#### 4.3 OPTIMIZATION OF MULTI-COUPLED CAVITY FILTERS

The application of multi-coupled cavity filters in modern microwave communication systems has received increasing attention. The theoretical work of Atia and Williams (1971, 1972) has inspired many advances in this area. These advances have been responsible for many improvements in satellite multiplexing networks, as has been discussed by Chen, Assal and Mahle (1976), Cameron (1982), Kudsia (1982), among others.

The growing variety and complexity of this type of filter necessitate the employment of modern CAD techniques. For example, the traditional approach to an analytical solution may become inappropriate when asynchronously tuned or nonminimum phase designs have to be considered. Bandler, Chen and Daijavad (1986a) have presented an efficient approach to the exact sensitivity analysis

of multi-coupled cavity filters, which has facilitated effective, flexible and systematic design optimization of these filters.

Beside their own significance, the filter examples in this section serve as illustrations of practical minimax nominal design optimization.

#### 4.3.1 The Physical Structure and the Equivalent Circuit

The typical structures for longitudinal dual-mode coupled cavity filters are shown in Fig. 4.4. The physical parameters include the geometrical dimensions of the cross slots through which the cavities are coupled and the penetrations of the coupling screws by which different modes in the same physical cavity are coupled. The cavity resonant frequencies may also be adjusted using the tuning screws.

The narrowband unterminated equivalent circuit introduced by Atia and Williams (1971) is described by a symmetrical impedance matrix as

$$\mathbf{Z} = j(s\mathbf{1} + \mathbf{M}) + r\mathbf{1}, \quad (4.8)$$

where  $\mathbf{1}$  denotes an  $N$  by  $N$  identity matrix and  $s$  is the normalized frequency variable given by

$$s = \frac{\omega_0}{\Delta\omega} \left( \frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right), \quad (4.9)$$

$\omega_0$  and  $\Delta\omega$  being the synchronously tuned cavity resonant frequency and the bandwidth parameter, respectively. In (4.8),  $r$  assumes the uniform cavity dissipation which is zero for a lossless filter.  $r$  is related to the unloaded  $Q$ -factor by

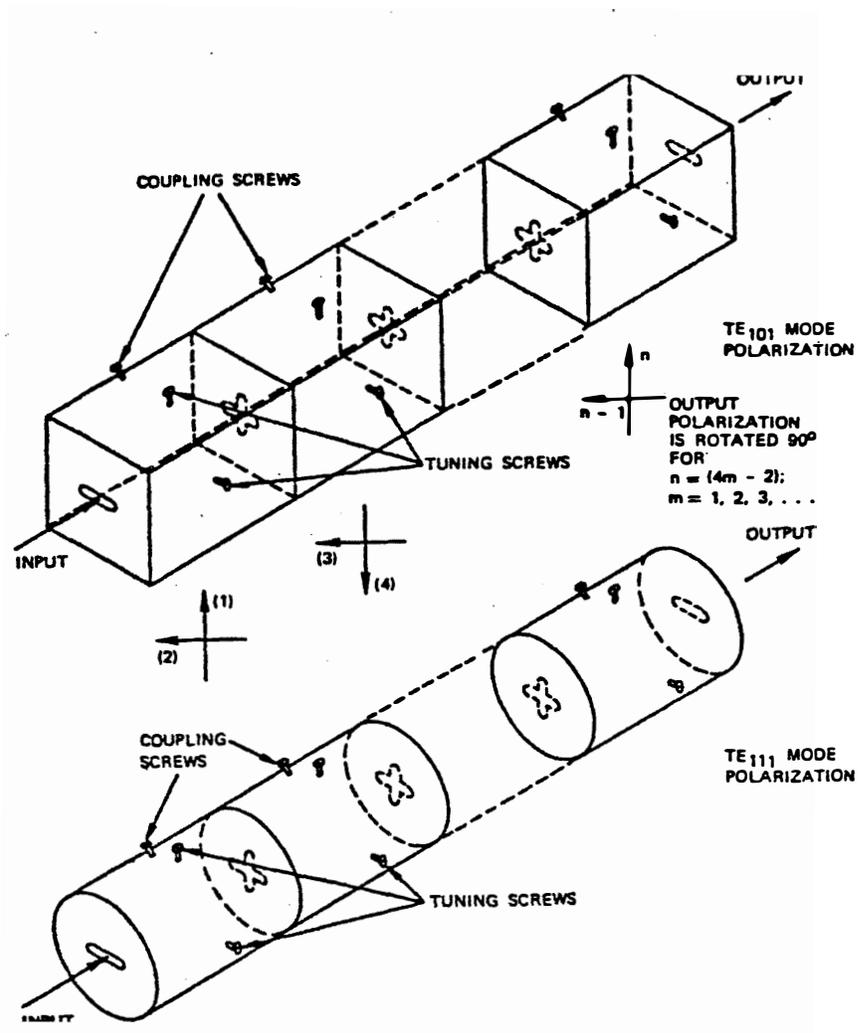


Fig. 4.4 Typical structures of longitudinal dual-mode cavity filters (Atia and Williams 1971).

$$r = \omega_0 / (\Delta\omega Q). \quad (4.10)$$

$\mathbf{M}$  is an  $N$  by  $N$  coupling matrix whose element  $M_{ij}$  represents the normalized coupling between the  $i$ th and  $j$ th cavities. The diagonal entry  $M_{ii}$  represents a deviation from synchronous tuning.  $\mathbf{M}$  is a symmetrical matrix, i.e.,  $M_{ij} = M_{ji}$ . Not all the elements in  $\mathbf{M}$  correspond to desirable and designable couplings. Some of them may indicate stray couplings. Dispersion effects on the filter can be modeled by a frequency dependent  $\mathbf{M}$  matrix. Fig. 4.5 depicts the equivalent circuit.

#### 4.3.2 Efficient Simulation and Sensitivity Calculations

In Chapter 3, we have described an efficient approach to sensitivity analysis of general networks. It can be applied most beneficially to the case of multi-coupled cavity filters.

From the nodal description  $\mathbf{Z} \mathbf{I} = \mathbf{V}$ , where  $\mathbf{Z}$  is given by (4.8), we define and solve

$$\mathbf{Z} \mathbf{p} = \mathbf{u}_1, \quad (4.11)$$

$$\mathbf{Z} \mathbf{q} = \mathbf{u}_N,$$

where  $\mathbf{u}_1 = [1 \ 0 \ \dots \ 0]^T$  and  $\mathbf{u}_N = [0 \ \dots \ 0 \ 1]^T$ . The unterminated filter is then modeled by a two-port whose short-circuit admittance matrix is given by

$$\mathbf{y} = \begin{bmatrix} Y_{11} & Y_{1N} \\ Y_{N1} & Y_{NN} \end{bmatrix} = \begin{bmatrix} P_1 & Q_1 \\ P_N & Q_N \end{bmatrix}. \quad (4.12)$$

Following the results of Section 3.4.1, the sensitivities of the  $\mathbf{y}$  matrix are obtained as

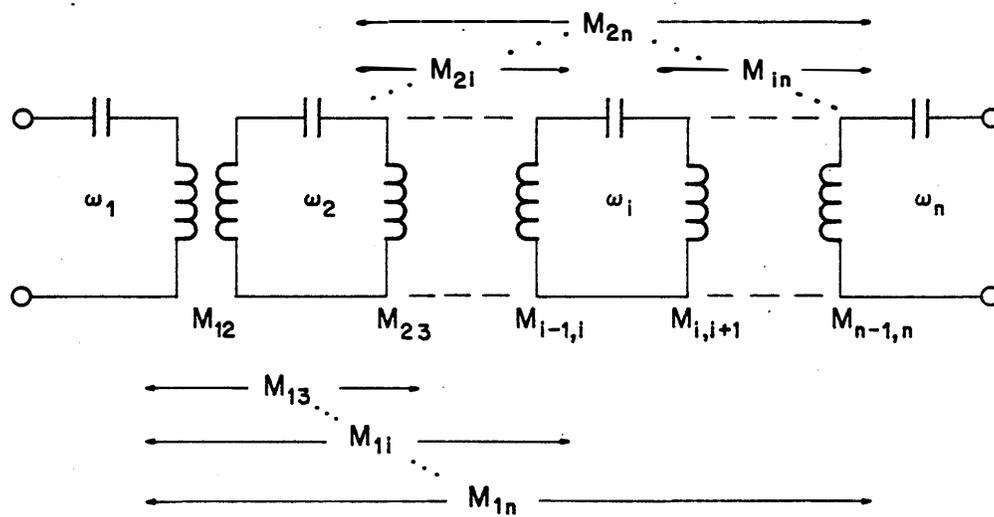


Fig. 4.5 Lumped-element equivalent circuit for an unterminated multi-coupled cavity filter.

$$\frac{\partial \mathbf{y}}{\partial \phi} = - \sum_{i=1}^N \sum_{k=1}^N \frac{\partial Z_{ik}}{\partial \phi} \begin{bmatrix} p_i p_k & p_i q_k \\ q_i p_k & q_i q_k \end{bmatrix}. \quad (4.13)$$

Since  $Z$  is symmetrical, we need only two solutions of the circuit (a nonreciprocal two-port would require four solutions).

For instance, consider  $\phi = M_{ik}$ ,

$$\frac{\partial \mathbf{y}}{\partial M_{ik}} = - c_j \begin{bmatrix} 2p_i p_k & (p_i q_k + p_k q_i) \\ (p_i q_k + p_k q_i) & 2q_i q_k \end{bmatrix}, \quad (4.14)$$

where  $c = 1$  if  $i \neq k$  (in which case  $M_{ik}$  appears in two places of  $Z$ , namely the  $(i,k)$  and  $(k,i)$  positions) or  $c = 1/2$  if  $i = k$ . Also,

$$\frac{\partial \mathbf{y}}{\partial \mathbf{r}} = - \begin{bmatrix} \mathbf{p}^T \mathbf{p} & \mathbf{p}^T \mathbf{q} \\ \mathbf{p}^T \mathbf{q} & \mathbf{q}^T \mathbf{q} \end{bmatrix}, \quad (4.15)$$

which is useful in predicting a first-order change due to cavity dissipation. We also have

$$\frac{\partial \mathbf{y}}{\partial \omega} = j s_\omega \frac{\partial \mathbf{y}}{\partial \mathbf{r}}, \quad (4.16)$$

where

$$s_\omega = \frac{\partial s}{\partial \omega} = \frac{1}{\Delta \omega} \left[ 1 + \left( \frac{\omega_0}{\omega} \right)^2 \right]. \quad (4.17)$$

We can use (4.16) in group delay and gain slope calculations.

By defining two additional solutions of the system as

$$\mathbf{Z} \bar{\mathbf{p}} = j s_\omega \mathbf{p}, \quad (4.18)$$

$$\mathbf{Z} \bar{\mathbf{q}} = j s_\omega \mathbf{q},$$

and following (3.25), we can evaluate

$$\frac{\partial}{\partial M_{ik}} \left( \frac{\partial \mathbf{y}}{\partial \omega} \right) = 2 \begin{bmatrix} \mathbf{p}^T \\ \mathbf{q}^T \end{bmatrix} \frac{\partial \mathbf{Z}}{\partial M_{ik}} [\bar{\mathbf{p}} \quad \bar{\mathbf{q}}]. \quad (4.19)$$

These second-order sensitivities become very useful in group delay

optimization, as illustrated later in this chapter.

We enjoy a computational advantage in the analysis of lossless filters. Since  $r = 0$  the impedance matrix  $Z$  in (4.8) is purely imaginary. Consequently, the systems defined by (4.11) and (4.18) can be solved by real arithmetic, which is obviously less time consuming than complex calculations.

An interesting special case is the canonical symmetrical filter structure (Kudisia, 1982). The coupling matrix of such a filter exhibits a dual symmetry with respect to its anti-diagonal as well as its diagonal, meaning that  $M_{ab} = M_{ik}$ , for  $a = N + 1 - k$  and  $b = N + 1 - i$ . Using a matrix notation, it implies

$$\begin{aligned} \mathbf{M} &= \bar{\mathbf{l}}\mathbf{M}\bar{\mathbf{l}}, \\ \mathbf{Z} &= \bar{\mathbf{l}}\mathbf{Z}\bar{\mathbf{l}}, \end{aligned} \tag{4.20}$$

where  $\bar{\mathbf{l}}$  is rotation matrix which has 1's on its anti-diagonal and zeros elsewhere. It can be easily verified that  $\bar{\mathbf{l}}\bar{\mathbf{l}} = \mathbf{1}$ .

Comparing

$$\mathbf{Z} \bar{\mathbf{l}}\mathbf{p} = \bar{\mathbf{l}}\mathbf{Z}\bar{\mathbf{l}} \bar{\mathbf{l}}\mathbf{p} = \bar{\mathbf{l}} \mathbf{Z}\mathbf{p} = \bar{\mathbf{l}} \mathbf{u}_1 = \mathbf{u}_N \tag{4.21}$$

with (4.11), we find that

$$\mathbf{q} = \bar{\mathbf{l}} \mathbf{p} , \tag{4.22}$$

i.e.,  $q_i = p_{N+1-i}$ . Similarly,  $\bar{\mathbf{q}} = \bar{\mathbf{l}} \bar{\mathbf{p}}$ . In other words, for this type of filters, the solutions of only two systems, instead of four, are sufficient for the evaluation of the first- and second-order sensitivities.

Once the two-port analysis is completed, various frequency responses of the filter and their sensitivities can be treated using the general formulas given in Table 3.1.

### 4.3.3 Cubic Interpolation in Minimax Design

In order to apply mathematical optimization techniques to a filter design problem, discrete frequency samples have to be considered. In minimax design, a poor selection of frequency samples may cause some difficulties, especially for a high-order Chebyshev filter whose responses exhibit many ripples. If the peaks of some ripples are missed (in frequency sampling), then the discretized solution may not be adequately close to the continuous minimax optimum. Conventionally, we try to overcome this difficulty by using densely spaced sample points. This, however, may lead to a prohibitively large number of error functions to be minimized.

Bandler and Chen (1984a) employed a cubic interpolation technique to detect the ripples of the responses and keep track of their locations during the optimization process. Consequently the frequency samples can be automatically and optimally selected.

Let  $e$  be a function which is continuous and differentiable in  $\omega$ . A ripple peak of  $e$  is a local maximum with respect to  $\omega$  and characterized by  $e' = \partial e / \partial \omega = 0$  and  $\partial^2 e / \partial \omega^2 < 0$ . This implies a change in the sign of  $\partial e / \partial \omega$  in the neighborhood of the maximum. It follows that if there exist two frequencies  $\omega_1 < \omega_2$  such that  $e'(\omega_1) > 0$  and  $e'(\omega_2) < 0$ , at least one maximum of  $e$  lies in between. If  $\omega_1$  and  $\omega_2$  are close enough to exclude the existence of multiple maxima, the cubic interpolation formula (Fletcher and Powell 1963) can be used to estimate the detected maximum as

$$\omega_m = \omega_2 - \frac{(\omega_2 - \omega_1)[x - y - e'(\omega_2)]}{e'(\omega_1) - e'(\omega_2) + 2x},$$

$$y = -e'(\omega_1) - e'(\omega_2) + 3 \frac{e(\omega_2) - e(\omega_1)}{\omega_2 - \omega_1}, \quad (4.23)$$

$$x = [y^2 - e'(\omega_1)e'(\omega_2)]^{\frac{1}{2}}.$$

This technique has proved to be very effective in practice. The response functions of interest as well as their sensitivities with respect  $\omega$  are evaluated at some base frequencies suitably selected (e.g., uniformly spaced points with adequate density). The ripples are then detected, located using (4.23) and chosen to form error functions to be optimized. Naturally, the number of such ripples are much smaller than the number of base points. In this way, the dimensionality of the optimization problem can be substantially reduced.

We have incorporated such techniques in the filter design examples of the following sections and, as a result, have been able to use a relatively small number of frequency samples to achieve virtually continuous minimax solutions.

#### 4.3.4 10th-Order Elliptic and Quasi-Elliptic

##### Self-Equalized Filters

A 10th-order multi-coupled cavity filter with a center frequency of 4GHz and a bandwidth of 40MHz is considered. It has a dual-symmetrical coupling matrix with 26 nonzero elements, namely  $M_{i,i+1}$ ,  $M_{i+1,i}$ ,  $i = 1, 2, \dots, 9$ , and  $M_{i,11-i}$ ,  $M_{11-i,i}$ ,  $i = 1, 2, 3, 4$ . Taking the dual-symmetry into account, we have 9 independently

designable couplings. The input and output transformer ratios are also considered as variables. The filter is assumed lossless.

The first example (Bandler and Chen 1984b) is a conventional elliptic (Chebyshev) bandpass filter. The design goal consists of an upper specification of 30dB on the return loss for the passband (3980MHz - 4020MHz) and a lower specification of 70dB on the insertion loss for the stopband (below 3976MHz or above 4024MHz). Since the amplitude responses are known to be symmetrical with respect to the center frequency, we need only consider one half of the operating frequency range. The cubic interpolation technique is employed to automatically determine, at each iteration, the positions of the frequency samples. The minimax solution is given in Table 4.1 and the filter responses are shown in Fig. 4.6.

Optimal trade-offs between the attenuation and group delay characteristics are required for high fidelity signal transmission. As has been demonstrated by Atia and Williams (1974), nonminimum-phase filters have the potential of realizing optimum amplitude and flat group delay characteristics.

Our second example is a 10th-order quasi-elliptic self-equalized filter achieved through simultaneous optimization of the amplitude and group delay (Bandler, Chen and Daijavad 1986a). Compared to the first example, the amplitude specifications are relaxed to be 22dB return loss for the passband and 45dB insertion loss for the stopband. An additional specification of 1.5ns is imposed on the relative group delay (delay variation) at four fixed sample points in the lower half of the passband, namely 3985MHz,

TABLE 4.1  
PARAMETERS FOR THE 10TH-ORDER ELLIPTIC FILTER

Parameter	Solution
$M_{12}, M_{21}, M_{9,10}, M_{10,9}$	0.97284
$M_{23}, M_{32}, M_{89}, M_{98}$	0.63006
$M_{34}, M_{43}, M_{78}, M_{87}$	0.54981
$M_{45}, M_{54}, M_{67}, M_{76}$	0.39867
$M_{56}, M_{65}$	0.88914
$M_{1,10}, M_{10,1}$	0.00298
$M_{29}, M_{92}$	-0.02422
$M_{38}, M_{83}$	0.15196
$M_{47}, M_{74}$	-0.49440
$n_1, n_2$	1.15823

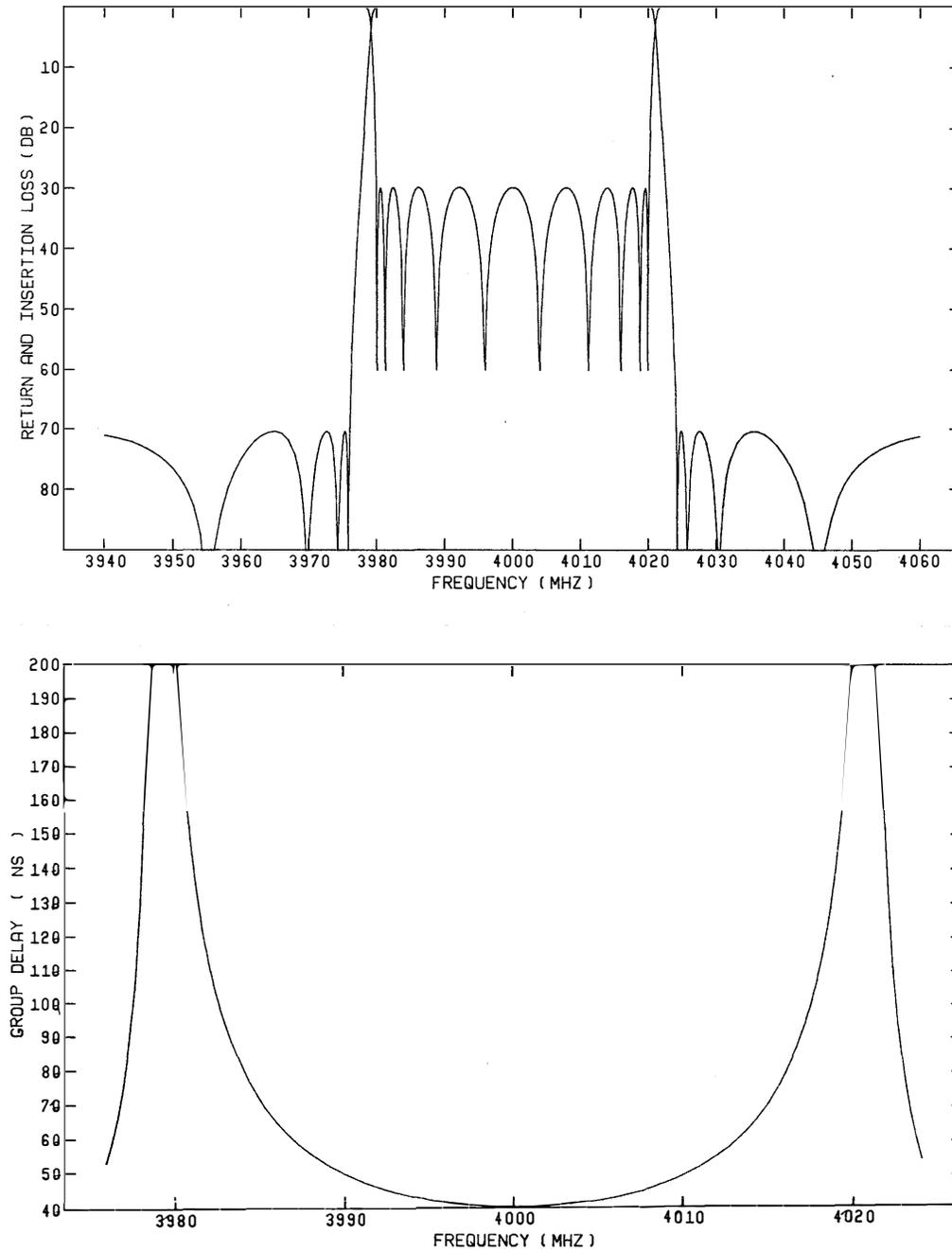


Fig. 4.6 Responses of the optimized 10th-order elliptic multi-coupled cavity filter.

3990MHz, 3995MHz and 4000MHz (since the passband group delay is relatively smooth, cubic interpolation was not used).

The solution is given in Table 4.2 and the responses are shown in Fig. 4.7. Although no explicit manipulation of the transfer function is necessary, we have achieved a nonminimum-phase design as expected. Fig. 4.8 shows that two zeros of the transfer function are located on the right hand half of the s-plane.

#### 4.3.5 An Asymmetric Design

The responses of a synchronously tuned filter, for which the nominal values of  $M_{1i}$  are zero, are always electrically symmetrical (with respect to the center frequency). Cameron (1982) has described the use of filters that have asymmetric characteristics, particularly in satellite communication systems. Advantages can be gained in applications to contiguous band multiplexers, where asynchronously tuned filters are utilized to annihilate inherent asymmetric distortions (such as a dispersive slope) by building compensating characteristics into the nominal design. They are also utilized to avoid the use of dummy channels by making the cut-off slope sharper on one side of the passband of the first and the last channel filters.

The conventional synthesis procedure is not applicable for asymmetric filters and a far more complicated procedure is needed (Cameron 1982). From a prescribed transfer function a prototype coupling matrix is constructed, and a sequence of rotational transformations using appropriate pivoting matrices is performed to

TABLE 4.2

PARAMETERS FOR THE 10TH-ORDER QUASI-ELLIPTIC SELF-EQUALIZED FILTER

Parameter	Solution
$M_{12}, M_{21}, M_{9,10}, M_{10,9}$	0.84424
$M_{23}, M_{32}, M_{89}, M_{98}$	0.59318
$M_{34}, M_{43}, M_{78}, M_{87}$	0.54438
$M_{45}, M_{54}, M_{67}, M_{76}$	0.53059
$M_{56}, M_{65}$	0.46916
$M_{1,10}, M_{10,1}$	0.01597
$M_{29}, M_{92}$	-0.02673
$M_{38}, M_{83}$	-0.05570
$M_{47}, M_{74}$	0.13067
$n_1, n_2$	1.02258

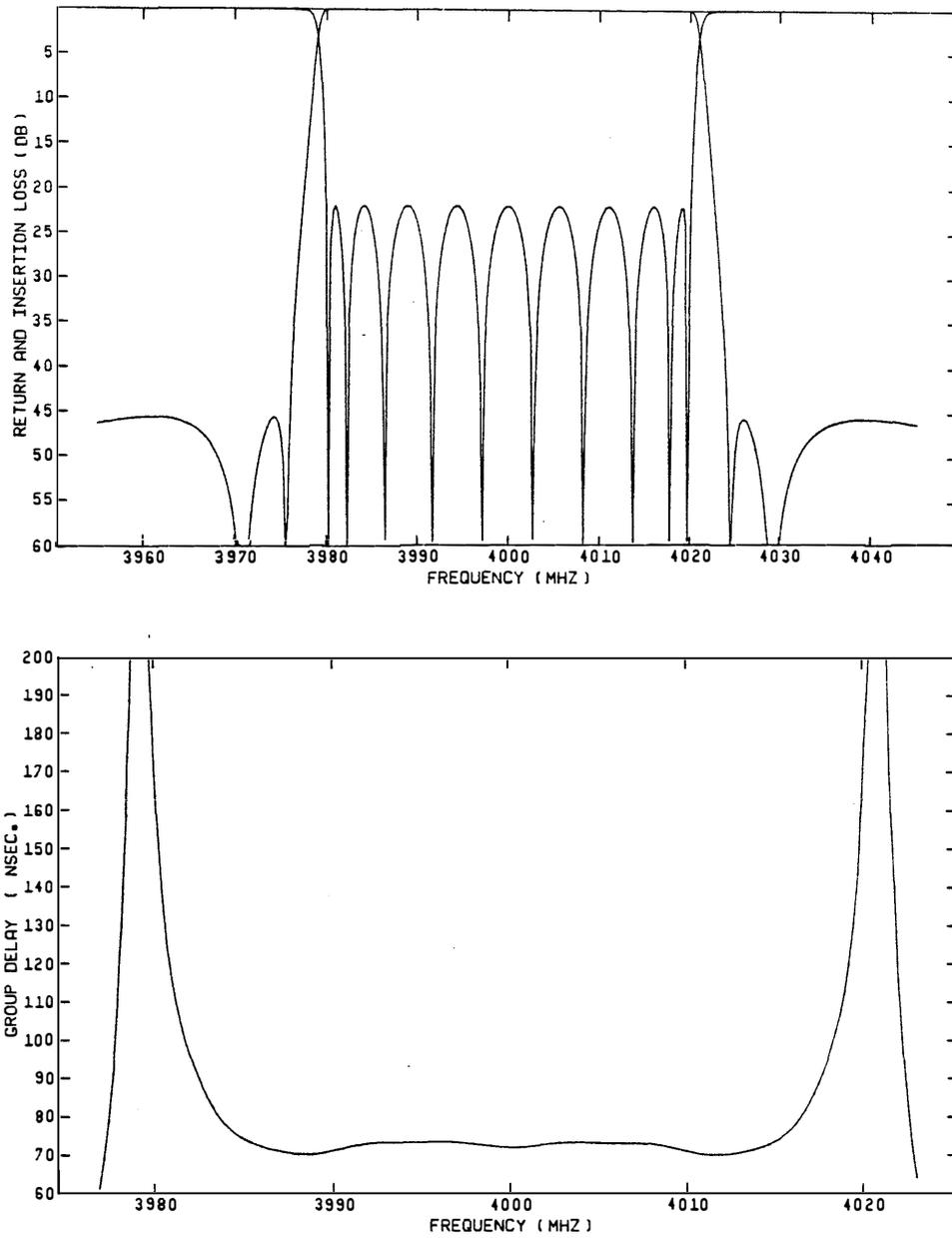


Fig. 4.7 Responses of the optimized 10th-order quasi-elliptic self-equalized filter.

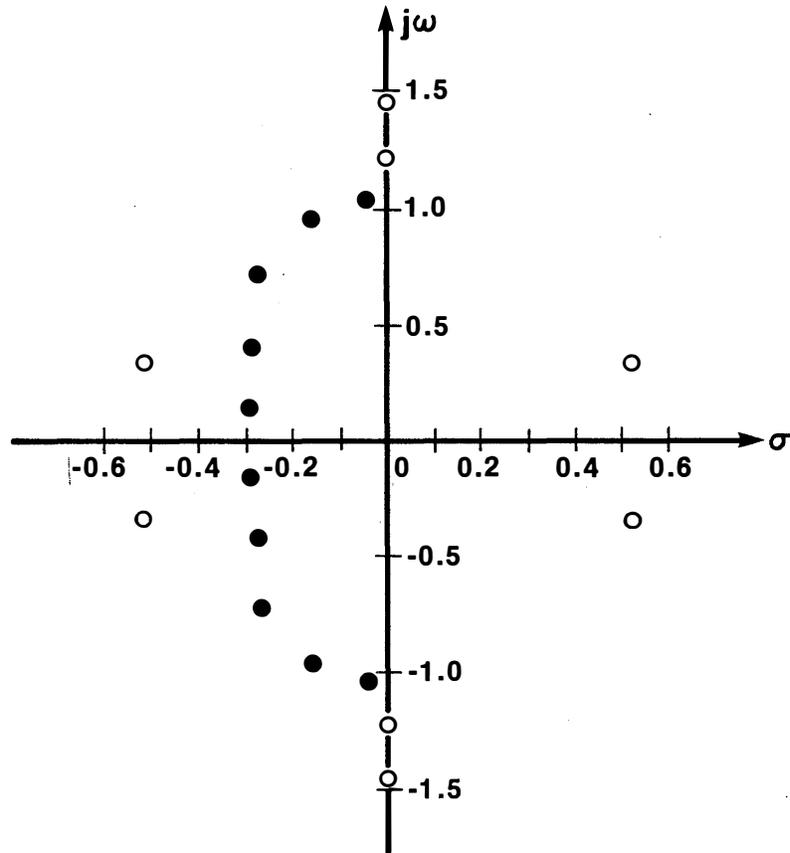


Fig. 4.8 The poles and zeros of the transfer function of the 10th-order self-equalized filter. • is a pole and o a zero. The zeros on the right-hand half of the complex plane reveals a nonminimum-phase transfer function.

remove the unrealizable couplings.

In comparison, the optimization approach is very flexible. The diagonal elements of the coupling matrix, which represent deviations from synchronous tuning, are included as variables. The specifications are simply defined to be asymmetrical as desired and can be modified conveniently.

We present a 6th-order example. The filter is centered at 4GHz with a 40MHz bandwidth. Including the diagonal elements there are 10 independently designable couplings. The passband upper specification is 25dB return loss. The stopband lower specification is 40dB insertion loss. The lower stopband is defined as below 3978MHz, allowing only 2MHz for the lower transition band, whereas the upper stopband begins at 4035MHz, 15MHz apart from the passband.

At the solution, the amplitude responses, as plotted in Fig. 4.9, exhibit the desired asymmetry in an optimal equal-ripple manner. A very sharp cut-off slope is achieved at the lower end of the passband. The filter parameters are given in Table 4.3.

#### 4.3.6 First-Order Prediction of the Effect of Cavity Dissipation

Lossless filters are often used as an ideal model to obtain nominal designs. In reality, the actual devices are subject to cavity dissipation. The performance of a lossy filter can, of course, be re-evaluated by exact simulations. This would require complex matrix analyses and, if different values of the Q-factor are considered, such complex analyses would have to be repeated for

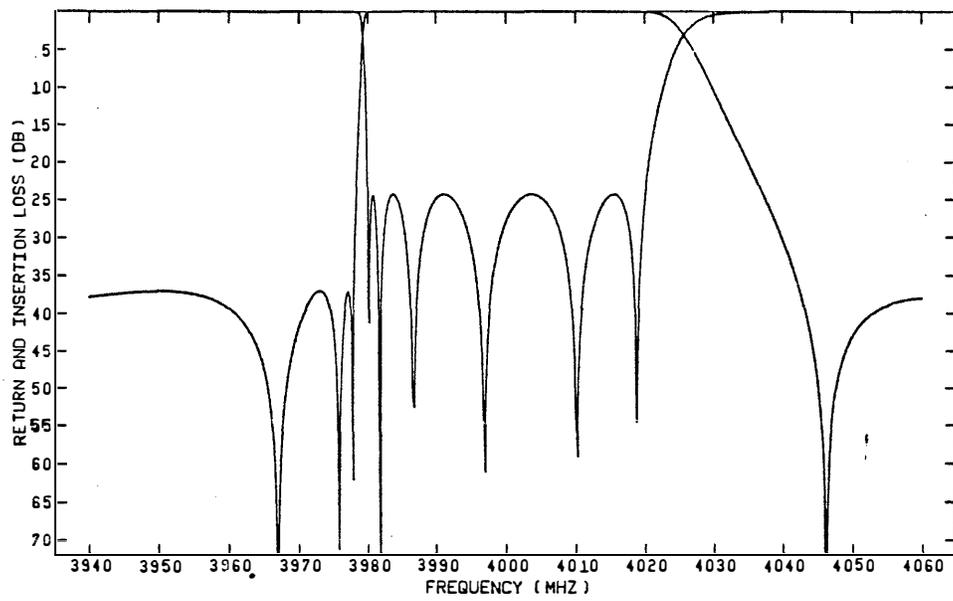


Fig. 4.9 Responses of the asymmetric 6th-order filter.

TABLE 4.3  
PARAMETERS FOR THE 6TH-ORDER ASYMMETRIC FILTER

Parameter	Solution
$M_{11}, M_{66}$	-0.05092
$M_{22}, M_{55}$	-0.07059
$M_{33}, M_{44}$	0.76097
$M_{12}, M_{21}, M_{56}, M_{65}$	0.91876
$M_{23}, M_{32}, M_{45}, M_{54}$	0.44780
$M_{34}, M_{43}$	0.24704
$M_{16}, M_{61}$	-0.08815
$M_{25}, M_{52}$	0.41092
$M_{15}, M_{51}, M_{26}, M_{62}$	0.01462
$M_{24}, M_{42}, M_{35}, M_{53}$	-0.31096
$n_1, n_2$	1.08666

each value of  $Q$ .

More efficiently, we can predict the non-ideal response by a first-order estimation. The basic sensitivity formula with respect to the cavity dissipation is given by (4.15). We have applied this method to one of our earlier examples, namely the 10th-order self-equalized filter. As shown in Fig. 4.10, the predicted passband insertion loss is indistinguishable from the exact simulation of the lossy filter (the numerical difference is less than 0.001dB).

It is a known fact that a filter with flat group delay characteristics is also less sensitive to dissipation. Bandler, Chen and Daijavad (1984b) have shown explicitly that the sensitivity of the amplitude with respect to dissipation is proportional to the group delay (this is also clear from (4.16)). For the example depicted in Fig. 4.10, the insertion loss variation is less than 0.1dB over 80 percent of the passband.

#### 4.4 LARGE-SCALE OPTIMIZATION OF MANIFOLD MULTIPLEXERS

The design of contiguous band manifold microwave multiplexers has been a problem of significant interest (Atia 1974, Chen, Assal and Mahle 1976, Bandler, Chen, Daijavad and Kellermann 1984, Chen 1985). It also presents a practical example for large-scale nominal circuit optimization.

A typical multiplexer structure is shown in Fig. 4.11. It consists of multi-coupled cavity channel filters distributed along a waveguide manifold. The interface between a channel filter and

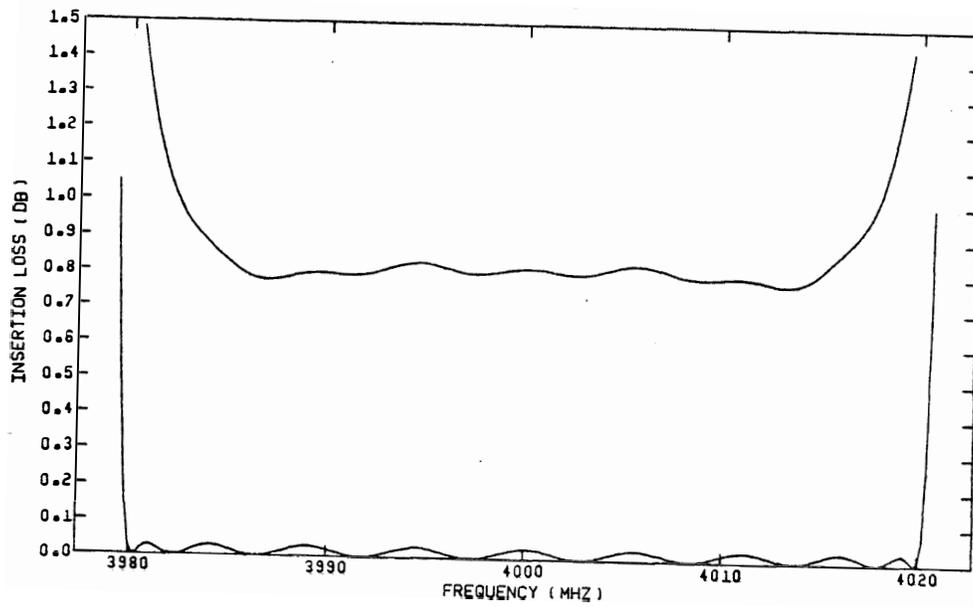


Fig. 4.10 Passband insertion loss of the 10th-order self-equalized filter for  $Q = 10,000$  and  $Q = \infty$ . The lossy response predicted by first-order sensitivity is indistinguishable from the result of exact simulation.

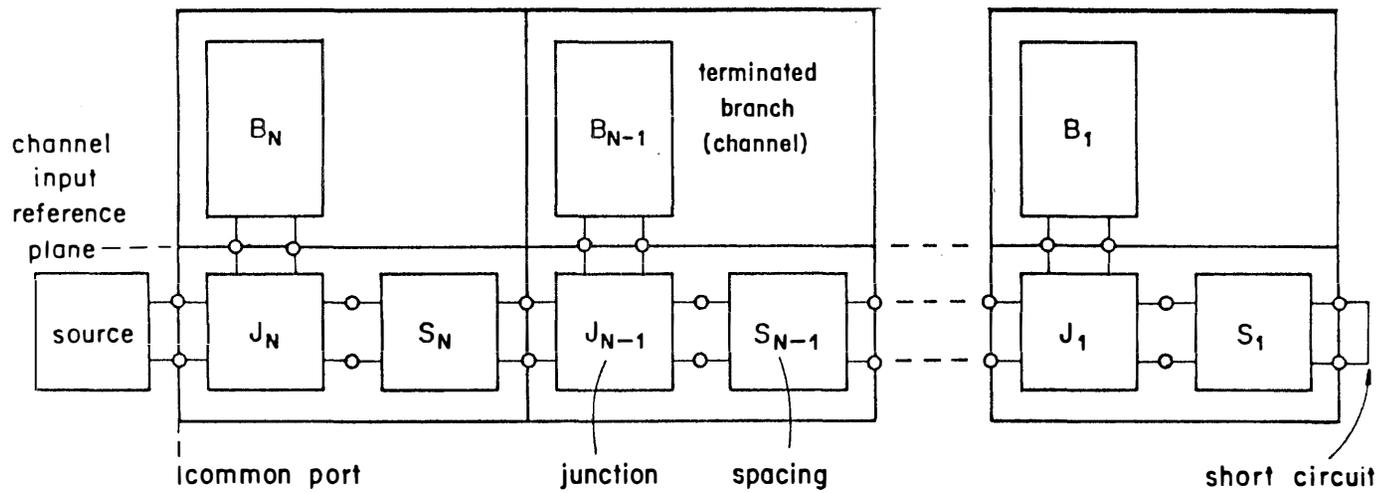


Fig. 4.11 A typical multiplexer structure.  $J_1, J_2, \dots, J_N$  are junction models,  $B_1, B_2, \dots, B_N$  are terminated channels and  $S_1, S_2, \dots, S_N$  represent spacings along the manifold.

the manifold waveguide can be a T-junction for which an empirical model has been proposed by Chen, Assal and Mahle (1976). The location of channel filters along the manifold can be modeled by spacings.

A major task in multiplexer optimization is to design the channel filters and to determine their spacing along the manifold. The responses of interest for a typical problem are common port return loss and insertion loss between the common port and each channel output port. Recently, a general approach to the simulation and sensitivity analysis of multiplexers has been presented by Bandler, Daijavad and Zhang (1986).

Bandler, Kellermann and Madsen (1987) have described a 12GHz 12-channel multiplexer which has a channel separation of 40MHz and a usable bandwidth of 39MHz. The center frequency of channel No. 1 is 12180MHz. Twelve 6th-order filters are used. A lower specification of 20 dB on the common port return loss is imposed over the passband of all 12 channels. The design process is started with a one-sided  $\ell_1$  optimization in order to deemphasize the worst violations of the specification and concentrate on the smaller errors. From the  $\ell_1$  solution minimax optimization is employed which involves 60 variables. The final optimized return loss is shown in Fig. 4.12.

Bandler, Chen, Daijavad, Kellermann, Renault and Zhang (1986) have described a novel decomposition approach to large-scale minimax optimization. It has been applied to expanding the 12-channel design into a 16-channel multiplexer which involves 240

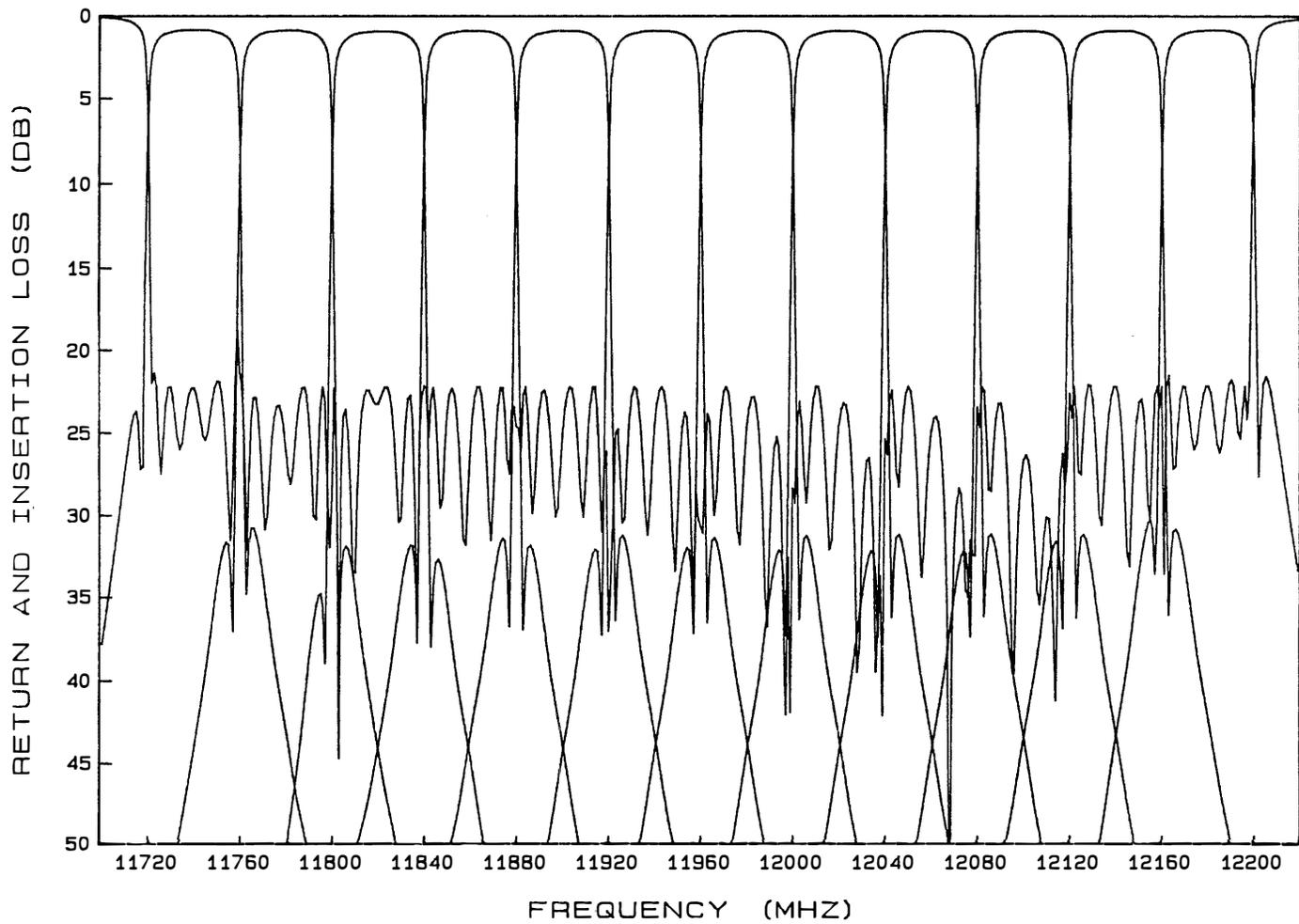


Fig. 4.12 Responses of the optimized 12 channel multiplexer (Bandler, Kellermann and Madsen 1987).

nonlinear designable parameters. Instead of making a blind attempt to optimize all the variables simultaneously, a suitable decomposition approach was taken in which we define, by adding one channel at a time, a sequence of localized problem involving a relatively small number of variables and functions. For example, when the 13th channel is added, we optimize only variables in channels 12 and 13 with specifications on responses in channels 11, 12 and 13. By repeating such a decomposition process four times, in which channels 13 to 16 are added and optimized successively, an optimal design is reached, as shown in Fig. 4.13.

Although the decomposition procedure may be justified intuitively, it is actually soundly based on sensitivity analyses. An automatic decomposition technique has been developed (Bandler and Zhang 1987) which can save computational time and alleviate memory storage problems for general large-scale applications.

#### 4.5 CONCLUDING REMARKS

In this chapter, we have demonstrated the application of optimization techniques to nominal circuit design. A hierarchy of simulation models has been introduced. Physical and model parameters and responses have been identified. A general formulation has been defined which leads to a nominal design through minimizing a suitable  $\ell_p$  measure of the errors between the given specifications and the calculated responses.

Illustrations of nominal design have been provided by the optimization of multi-coupled cavity filters as well as manifold

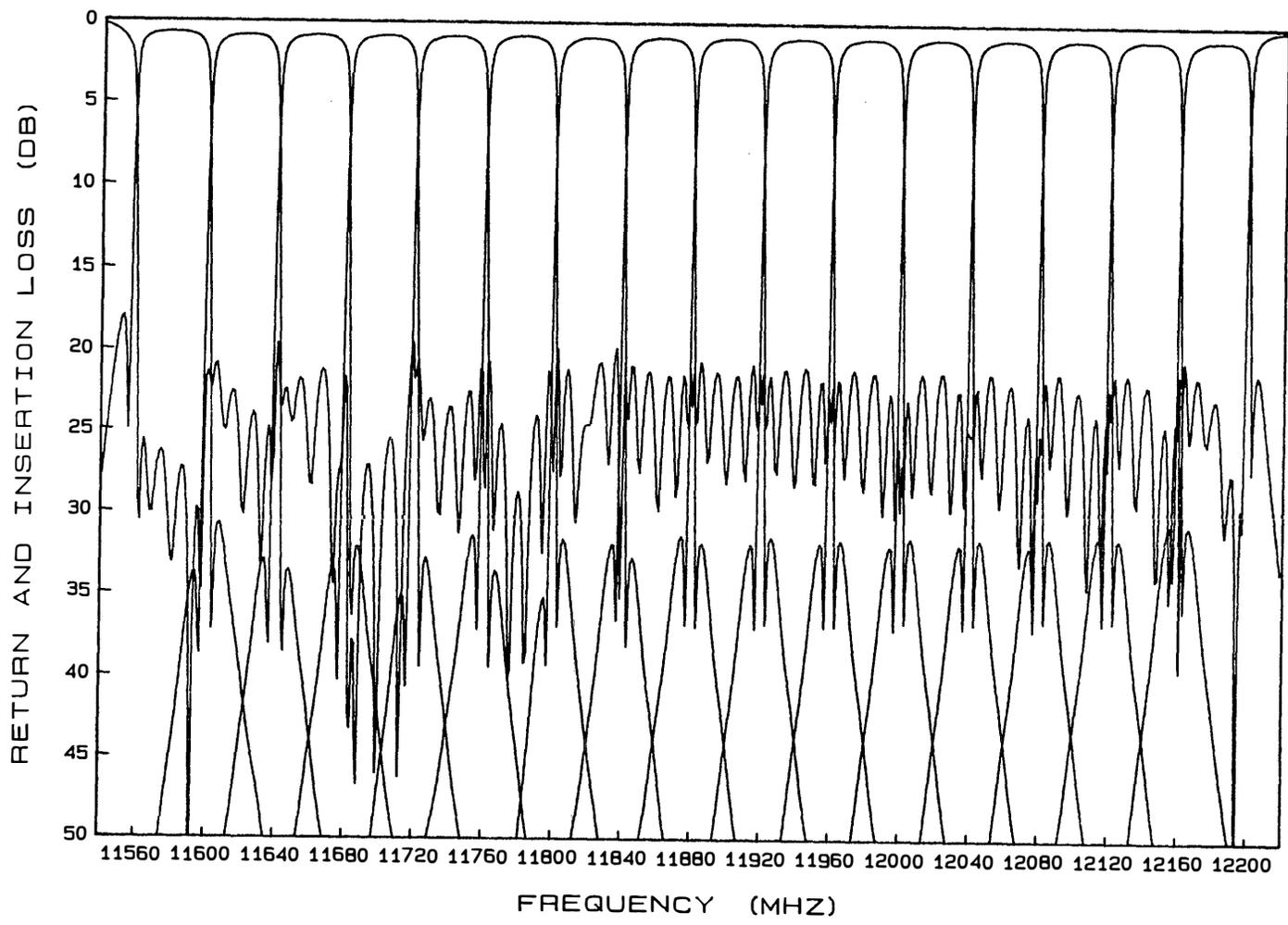


Fig. 4.13 Optimized responses of the 16 channel multiplexer.

multiplexers. Practical examples of significant interest have been described in detail. The use of the cubic interpolation techniques in minimax optimization has also been discussed.

The nominal design approach is based on an ideal assumption that the simulation models are precise and accurate. The removal of this assumption leads to more realistic approaches to circuit optimization, both in design and modeling, which will be addressed in the next two chapters.

## CHAPTER 5

### REALISTIC APPROACHES TO CIRCUIT DESIGN

#### 5.1 INTRODUCTION

The approach of nominal circuit optimization, which we have covered in Chapter 4, focuses attention on an idealized situation in which the models are assumed to be precise and accurate. In reality, unfortunately, there are parameter tolerances and model uncertainties to be accounted for. Such realistic considerations arise from design problems in which a large volume of production is envisaged, e.g., integrated circuits.

Recognition that an actual realization of a nominal design is subject to fluctuation or deviation led, in the past, to the so-called sensitivity minimization approach (e.g., Schoeffler 1964, Laker, Ghausi and Kelly 1975). Employed by filter designers, the approach involves measures of performance sensitivity, typically first-order, and including it in the objective function.

The statistical design approach emerged from the pioneering work of Karafin (1971), Butler (1971), Pinel and Roberts (1972) and Elias (1975). It deals more explicitly with process imprecision, manufacturing tolerances, model uncertainties, and so on, with the aim of improving production yield and reducing cost. There has been an increasing interest in statistical design in recent years, perhaps due to the growth of large and integrated circuits.

In this chapter, we first identify the possible tolerances

and uncertainties associated with a typical physical system and its simulation models. Multiple circuits are defined to relate these uncertainties to a nominal point. The concepts of centering, tolerancing and tuning are discussed in relation to yield enhancement and cost reduction.

A review of statistical design techniques is then given. We describe in some detail several representative methods including the worst-case design approach, the simplicial and multidimensional approximations, the gravity method and the parametric sampling method.

We also propose a generalized  $\ell_p$  centering algorithm as a natural extension to the  $\ell_p$  nominal optimization. It provides a unified formulation of yield enhancement for both the worst-case and the case where yield is less than 100%.

Circuit examples are also presented as illustrations.

## 5.2 A MULTI-CIRCUIT FORMULATION

### 5.2.1 Physical Tolerances and Model Uncertainties

Tolerances and uncertainties can be defined for the physical system and represented in the simulation models of different levels. For the physical system, consider

$$\begin{aligned} \mathbf{F}^M &= \mathbf{F}^{M,0}(\boldsymbol{\phi}^M) + \Delta\mathbf{F}^M, \\ \boldsymbol{\phi}^M &= \boldsymbol{\phi}^{M,0} + \Delta\boldsymbol{\phi}^M, \end{aligned} \tag{5.1}$$

where  $\Delta\mathbf{F}^M$  represents measurement errors,  $\boldsymbol{\phi}^{M,0}$  a nominal value for  $\boldsymbol{\phi}^M$  and  $\Delta\boldsymbol{\phi}^M$  some physical (manufacturing, operating) tolerances.

In order to represent the physical system more realistical-

ly, the hierarchy of simulation models, which we have introduced in Section 4.2.1, needs to be modified as

$$\begin{aligned}
 \mathbf{F}^{\text{H}} &= \mathbf{F}^{\text{H},0}(\mathbf{F}^{\text{L}}) + \Delta\mathbf{F}^{\text{H}}, \\
 \mathbf{F}^{\text{L}} &= \mathbf{F}^{\text{L},0}(\phi^{\text{H}}) + \Delta\mathbf{F}^{\text{L}}, \\
 \phi^{\text{H}} &= \phi^{\text{H},0}(\phi^{\text{J}}) + \Delta\phi^{\text{H}}, \\
 \phi^{\text{J}} &= \phi^{\text{J},0} + \Delta\phi^{\text{J}},
 \end{aligned} \tag{5.2}$$

where  $\phi^{\text{J},0}$ ,  $\phi^{\text{H},0}$ ,  $\mathbf{F}^{\text{L},0}$  and  $\mathbf{F}^{\text{H},0}$  are nominal models applicable at different levels.  $\Delta\phi^{\text{J}}$ ,  $\Delta\phi^{\text{H}}$ ,  $\Delta\mathbf{F}^{\text{L}}$  and  $\Delta\mathbf{F}^{\text{H}}$  represent uncertainties or inaccuracies associated with the respective models.  $\Delta\phi^{\text{J}}$  corresponds to the tolerances  $\Delta\phi^{\text{M}}$ .  $\Delta\phi^{\text{H}}$  may be due to the approximate nature of an empirical formula. Parasitic effects which are not adequately modeled in  $\phi^{\text{H}}$  will contribute to  $\Delta\mathbf{F}^{\text{L}}$ , and finally we attribute anything else that causes a mismatch between  $\mathbf{F}^{\text{H},0}$  and  $\mathbf{F}^{\text{M},0}$  to  $\Delta\mathbf{F}^{\text{H}}$ .

These concepts can be illustrated by the one-section strip-line transformer example due to Bandler, Liu and Tromp (1976b) which has also been considered in Section 4.2.1. Tolerances may be imposed on the physical parameters including the strip widths and thicknesses, the dielectric constants, the section length and substrate thicknesses (see Fig. 4.1). Such tolerances correspond to  $\Delta\phi^{\text{M}}$  and are represented in the model by  $\Delta\phi^{\text{J}}$ . We may use  $\Delta\phi^{\text{H}}$  to represent uncertainties associated with the empirical formulas which relate the physical parameters to the equivalent circuit parameters (the effective linewidths, the junction inductances and the effective section length). Mismatches in the terminations at different frequencies may be estimated by  $\Delta\mathbf{F}^{\text{H}}$ , which contribute to

the discrepancy between the actual and the nominal (assuming matched terminations) reflection coefficients.

The distinction between different levels of model uncertainties can be quite subtle. For example, consider the parasitic resistance  $r$  associated with an inductor whose inductance is  $L$ . Both  $L$  and  $r$  are functions of the number of turns of a coil (which is a physical parameter). Depending on whether or not  $r$  has been modeled by the equivalent circuit (i.e., whether or not  $r$  has been included in  $\phi^H$ ), the uncertainty associated with  $r$  may appear in  $\Delta\phi^H$  or in  $\Delta F^L$ .

### 5.2.2 Multiple Circuits and Yield

If our primary concern is to improve production yield and reduce cost in the presence of the tolerances  $\Delta\phi^L$  and the model uncertainties  $\Delta\phi^H$ , a single-circuit nominal design will not be adequate. We have to consider, as representatives of the actual production outcomes, multiple circuits defined by

$$\phi^k = \phi^0 + s^k \quad k = 1, 2, \dots, K, \quad (5.3)$$

where  $\phi^0$ ,  $\phi^k$  and  $s^k$  are generic notation for the nominal parameters, the  $k$ th set of parameters and a deviate due to the uncertainties, respectively. A more elaborate definition will be given as we proceed.

For each circuit, we define an acceptance index by

$$I_a(\phi) = \begin{cases} 1 & \text{if } H(\mathbf{e}(\phi)) \leq 0 \\ 0 & \text{otherwise} \end{cases}, \quad (5.4)$$

where  $H(\mathbf{e})$  is a generalized  $\ell_p$  function defined in Section 2.2.1,

and  $H(\mathbf{e}(\boldsymbol{\phi})) \leq 0$  indicates satisfaction of the specifications by  $\boldsymbol{\phi}$ . An estimate of the yield is given by the percentage of acceptable samples out of the total, as

$$Y \approx \left[ \sum_{k=1}^K I_a(\boldsymbol{\phi}^k) \right] / K . \quad (5.5)$$

The merit of a design can then be judged more realistically according to the yield it promises. Fig. 5.1 shows three nominal points and the related yield. Now we shall have a closer look at the definition of multiple circuits.

In the Monte Carlo method the deviates  $\mathbf{s}^k$  are constructed from random numbers generated by a physical process or arithmetical algorithms. Typically, we assume a statistical distribution for  $\Delta\boldsymbol{\phi}^L$ , denoted by  $D^L(\boldsymbol{\epsilon}^L)$  where  $\boldsymbol{\epsilon}^L$  is a vector of tolerance variables. For example, we may consider a multidimensional uniform distribution on  $[-\boldsymbol{\epsilon}^L, \boldsymbol{\epsilon}^L]$ . Similarly, we assume a  $D^H(\boldsymbol{\epsilon}^H)$  for  $\Delta\boldsymbol{\phi}^H$ .

At the low level, consider

$$\boldsymbol{\phi}^{L,k} = \boldsymbol{\phi}^{L,0} + \mathbf{s}^{L,k}, \quad k = 1, 2, \dots, K^L, \quad (5.6)$$

where  $\mathbf{s}^{L,k}$  are samples from the distribution  $D^L$ . At the higher level, we have, for each  $k$ ,

$$\boldsymbol{\phi}^{H,k,i} = \boldsymbol{\phi}^{H,0} + \mathbf{s}^{H,k,i}, \quad i = 1, 2, \dots, K^H, \quad (5.7)$$

where

$$\begin{aligned} \boldsymbol{\phi}^{H,0} &= \boldsymbol{\phi}^{H,0}(\boldsymbol{\phi}^{L,0}), \\ \mathbf{s}^{H,k,i} &= \boldsymbol{\phi}^{H,0}(\boldsymbol{\phi}^{L,k}) - \boldsymbol{\phi}^{H,0}(\boldsymbol{\phi}^{L,0}) + \boldsymbol{\delta}^{k,i}, \end{aligned} \quad (5.8)$$

with  $\boldsymbol{\delta}^{k,i}$  being samples from the distribution  $D^H$ .

One might propose a distribution for  $\mathbf{s}^{H,k,i}$  which would presumably encompass the effect of both  $D^L$  and  $D^H$ . But, while we

may reasonably assume simple and independent distributions for  $\Delta\phi^I$  and  $\Delta\phi^H$ , the compound distribution is likely to be complicated and correlated and, therefore, much less desirable.

### 5.2.3 Centering, Tolerancing and Tuning

Again, in order to simplify the notation, we use  $\phi^0$  for the nominal circuit and  $\epsilon$  for the tolerance variables.

Design centering is generally related to the enhancement of yield, directly or indirectly. An important class of problems involves design centering with fixed tolerances, usually relative to the corresponding nominal values. This is called the fixed tolerance problem (FTP). The optimization variables are elements of the nominal circuit parameters  $\phi^0$ . Incidentally, the nominal optimization problem (i.e., single-circuit design) is sometimes referred to as the zero tolerance problem (ZTP).

Since imposing tight tolerances on the circuit parameters will increase the cost of component fabrication or process operation, we may attempt to maximize the allowable tolerances subject to an acceptable yield. In this case both  $\phi^0$  and  $\epsilon$  are considered as variables to be optimized. Such a problem is referred to as optimal tolerancing, optimal tolerance assignment, or the variable tolerance problem (VTP).

Tuning some components of  $\phi^M$  after production, whether by the manufacturer or by a customer, is quite commonly used as a means of improving the yield. This process can also be simulated using the model by introducing a vector of designable tuning adjus-

tments  $\tau^k$  for each circuit, as

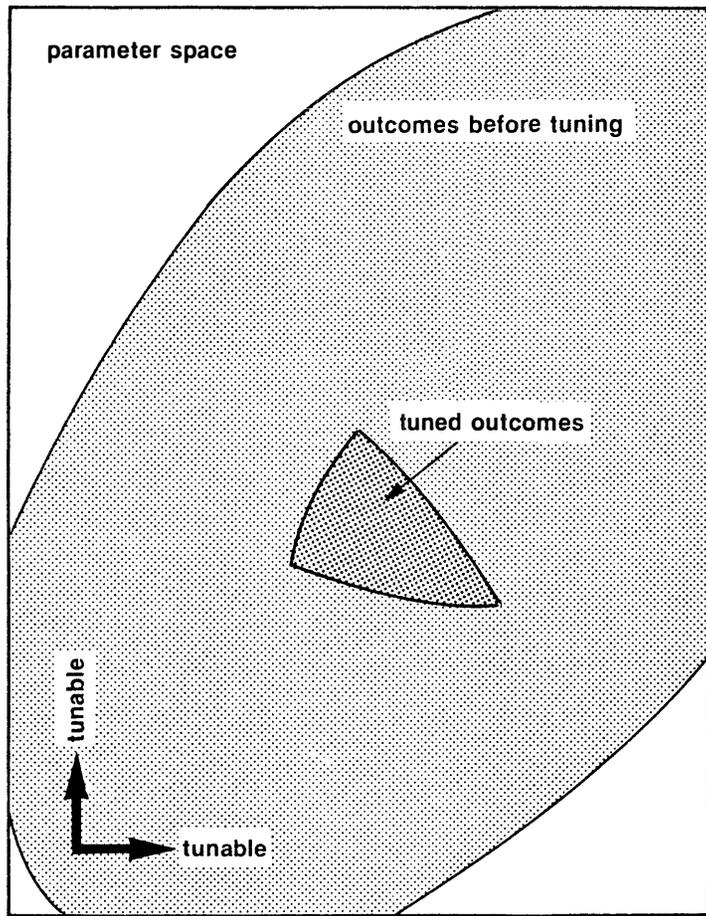
$$\phi^k = \phi^0 + s^k + \tau^k, \quad k = 1, 2, \dots, K. \quad (5.9)$$

We have to determine, through optimization, the value of  $\tau^k$  such that the specifications will be satisfied at  $\phi^k$  which may otherwise be unacceptable, as depicted in Figs. 5.2 and 5.3. The introduction of tuning, on the other hand, also increases design complexity and manufacturing cost. We seek a suitable compromise by solving an optimization problem in which  $\tau^k$  are treated as part of the variables. Analogously to ZTP, FTP and VTP we can define zero tuning, fixed tuning and variable tuning problems (Bandler and Kellermann 1983).

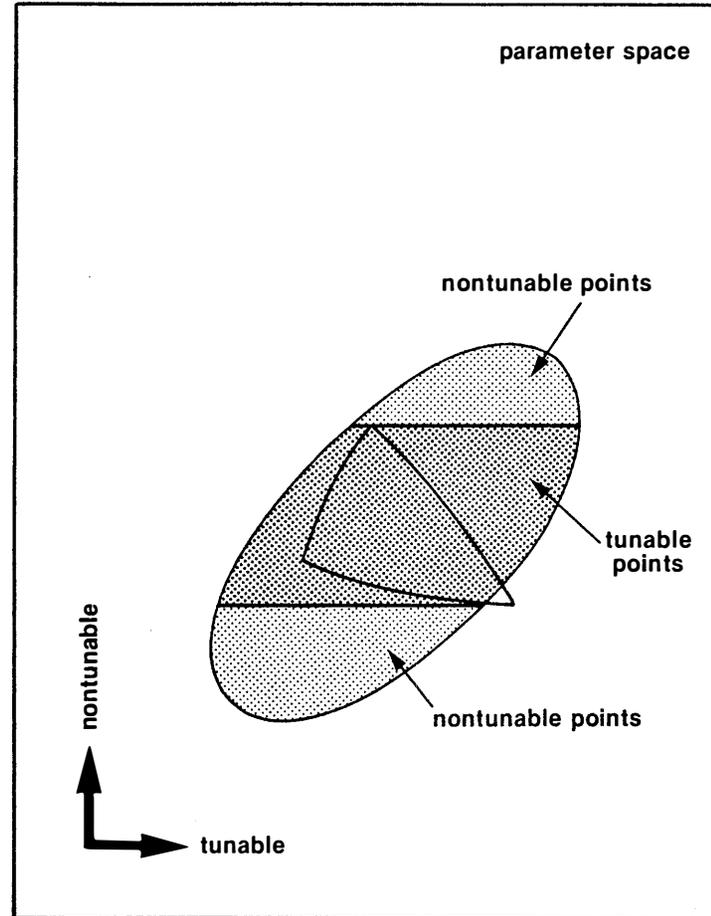
From nominal design, centering, optimal tolerancing to optimal tuning, we have defined a range of problems which lead to increasingly improved yield but, on the other hand, correspond to increasing complexity. An increase in yield does not necessarily lead to a decrease in cost. A commonly assumed cost versus yield curve (Singhal and Pinel 1981) is shown in Fig. 5.4. Often, a rather abstract objective function (cost function) is selected for the cost-yield design problem. Fig. 5.5 illustrates a design with a 100% yield and a second design corresponding to the minimum cost.

### 5.3 A REVIEW OF TECHNIQUES FOR STATISTICAL DESIGN

The statistical design approach was pioneered by Karafin (1971), Butler (1971), Pinel and Roberts (1972), Elias (1975), as well as Bandler, Liu and Tromp (1976a). During the years, significant contributions have been made by, among others, Director and



(a)



(b)

Fig. 5.2 Illustrations of tuning: (a) both parameters are tunable and (b) only one parameter is tunable. In case (a) the probability that an untuned design meets the specifications is very low.

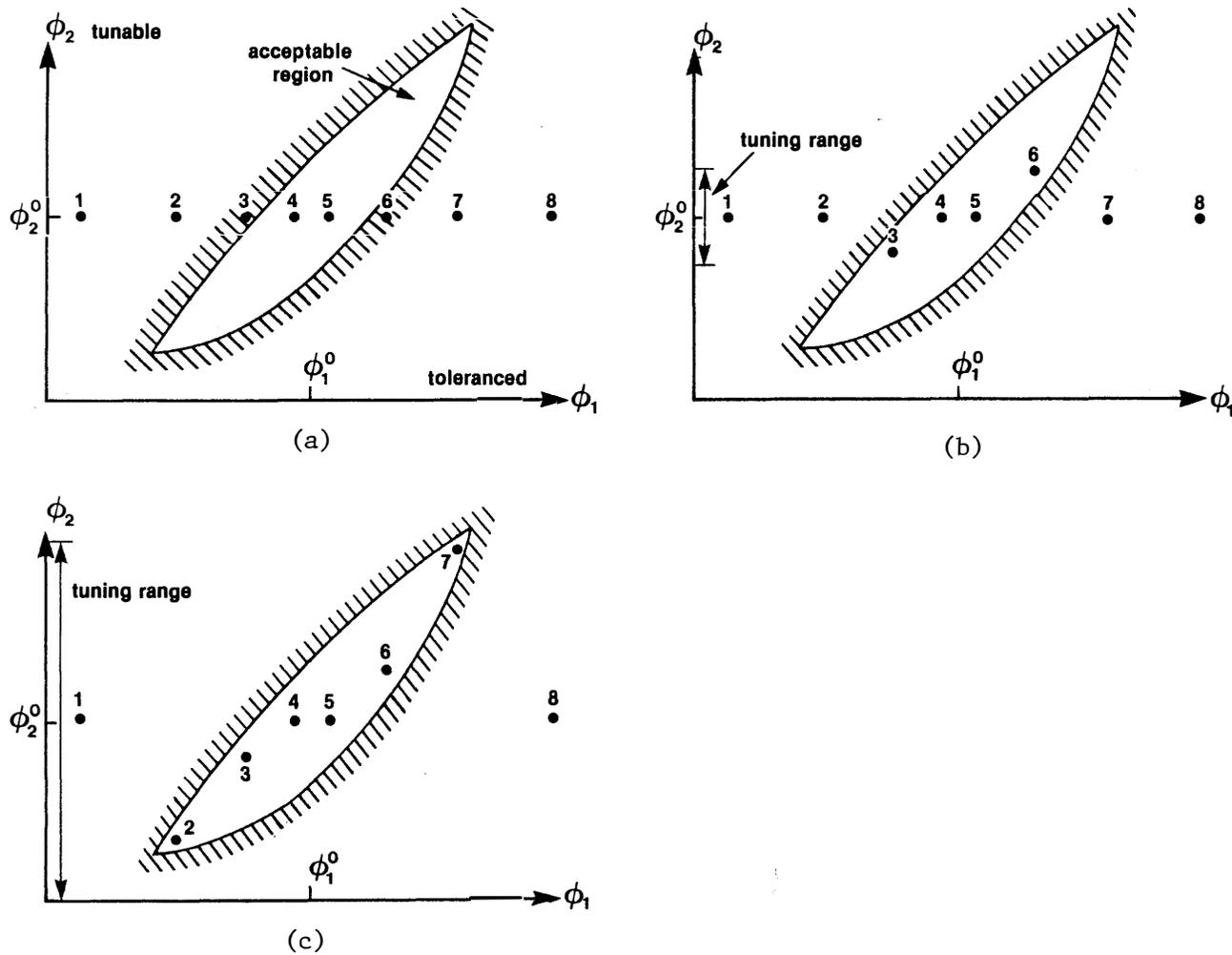


Fig. 5.3 Illustration of multi-circuit design considering eight outcomes.  $\phi_1$  is tolerated and  $\phi_2$  is tunable. (a) Without tuning the yield is  $2/8$  (25%). (b) Tuning on  $\phi_2$  is restricted to a small range. The improved yield is  $4/8$  (50%). (c) A 75% yield is achieved by allowing a large tuning range.

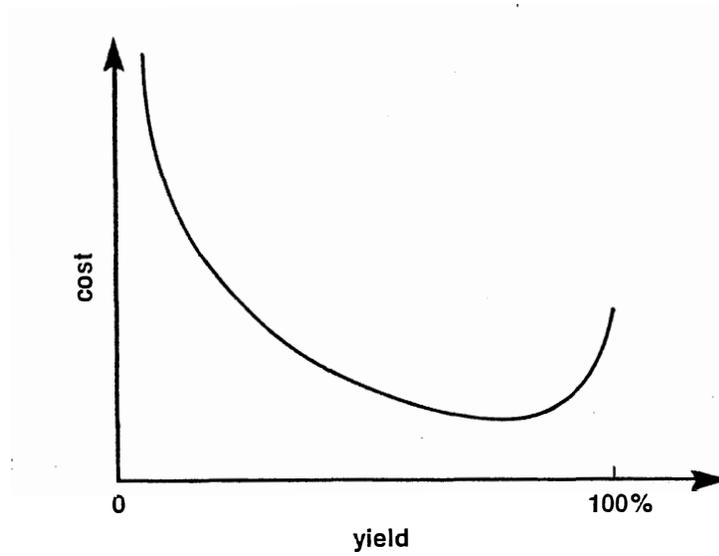


Fig. 5.4 A typical cost-versus-yield curve (Singhal and Pinel 1981).

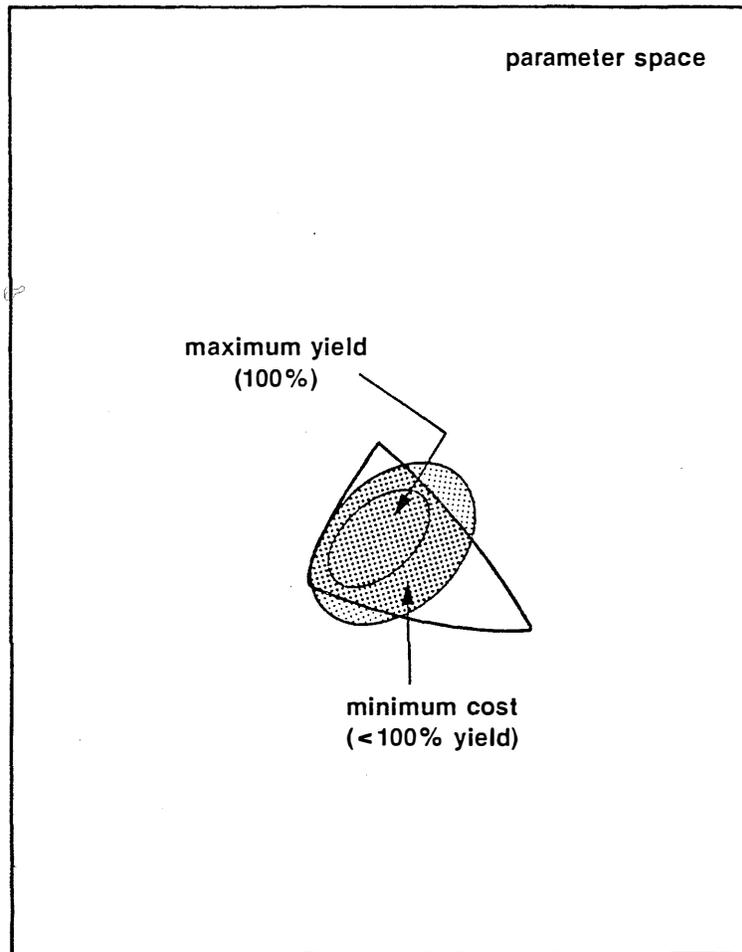


Fig. 5.5 A maximum yield design and a minimum cost design.

Hachtel (the simplicial method, 1977), Bandler and Abdel-Malek (multidimensional approximation, 1978), Polak and Sangiovanni-Vincentelli (a method using outer approximation, 1979), Soin and Spence (the gravity method, 1980), Singhal and Pinel (the parametric sampling method, 1981), as well as Biernacki and Styblinski (dynamic constraint approximation, 1986).

In this section, we describe in some detail several representative techniques for statistical design.

### 5.3.1 Worst-Case Design

At the heart of the worst-case approach is an attempt to achieve a 100% yield. Since this implies that the specifications have to be satisfied for all the possible outcomes, it will suffice to consider only the worst cases.

Bandler, Liu and Tromp (1976a) have formulated worst-case design as a nonlinear programming problem as

$$\begin{aligned} & \text{minimize } C(\mathbf{x}) \\ & \mathbf{x} \\ & \text{subject to } \mathbf{e}(\phi^k) \leq \mathbf{0}, \text{ for all } k, \end{aligned} \tag{5.10}$$

where  $C(\mathbf{x})$  is a suitable cost function and  $\phi^k$  are candidates for the worst case. For instance, we may have

$$C(\mathbf{x}) = \sum_{i \in I_\epsilon} \frac{a_i}{\epsilon_i} + \sum_{i \in I_t} b_i t_i, \tag{5.11}$$

where  $I_\epsilon$  and  $I_t$  are index sets identifying the tolerated and tunable parameters, respectively.  $\epsilon_i$  and  $t_i$  are the tolerance and the tuning range, respectively, associated with the  $i$ th parameter.  $a_i$  and  $b_i$  are nonnegative weights. A cost function can also be

defined for relative tolerances and tuning by including  $\phi_1^0$  into (5.11).

A critical part of this approach is the determination of the worst cases. Vertices of the tolerance region, for example, are possible candidates for the worst case by assuming one-dimensional convexity. The yield function does not appear explicitly in (5.10), instead, a 100% yield is implied by a feasible solution.

Bandler and Charalambous (1974) have derived a solution to (5.10) by minimax optimization. Polak and Sangiovanni-Vincentelli (1979) have proposed a different but equivalent formulation which involves a nondifferentiable optimization.

The worst-case approach is not always appropriate. While attempting to obtain a 100% yield, the solution may necessitate unrealistically tight tolerances or excessive tuning. In either case, the cost may be too high.

### 5.3.2 Approximations of the Acceptable Region

The acceptable region with respect to a given set of specifications is defined as

$$R_a = \{\phi \mid H(\mathbf{e}(\phi)) \leq 0\}, \quad (5.12)$$

where  $H(\mathbf{e})$  is an  $\ell_p$  function. Since the yield is given by the percentage of circuit outcomes that fall within the acceptable region, we may wish to find an approximation to that region.

Director and Hachtel (1977) have developed a simplicial approximation approach. It begins by determining some points  $\phi^k$  on the boundary of  $R_a$  which is given by  $\Omega_a = \{\phi \mid H(\mathbf{e}(\phi)) = 0\}$ . The

convex hull of these points forms a polyhedron. The largest hypersphere inscribed within the polyhedron gives an approximation to  $R_a$  and is found by solving a linear programming problem. Using line searches more points on the boundary are located and the polyhedron is expanded. The process thus provides a monotone increasing lower bound on the yield. The center and radius of the hypersphere can be used to determine the centered nominal point and the tolerances, respectively. The application of this method is, however, limited by the assumption of a convex acceptable region.

Bandler and Abdel-Malek (1978, 1980) presented a method which approximates each  $e_j(\phi)$  by a low-order multidimensional polynomial. Circuit simulations are performed at some  $\phi^k$  selected around a reference point. From the values of  $e_j(\phi^k)$  the coefficients of the approximating polynomial are determined by solving a linear system of equations. Suitable linear cuts are constructed to approximate the boundary  $\Omega_a$ . The yield is estimated through evaluation of the hypervolumes that lie outside  $R_a$  but inside the tolerance region. In critical regions these polynomial approximations may be subsequently updated during optimization. The one-dimensional convexity assumption for this method is much less restrictive than the multidimensional convexity required by the simplicial approach. Sensitivities for the estimated yield are also available.

Recently, Biernacki and Styblinski (1986) have extended the work on multidimensional polynomial approximation by considering a dynamic constraint approximation scheme. It avoids the large

number of base points required for a full quadratic interpolation by selecting a maximally flat interpolation. During optimization, whenever a new base point is added the approximation is updated. It leads to improved accuracy compared with a linear model and also reduced computational effort compared with a full quadratic model.

### 5.3.3 The Gravity Method

Soin and Spence (1980) proposed a statistical exploration approach. Based on a Monte Carlo analysis, the centers of gravity of the failed and passed samples are determined as, respectively,

$$\begin{aligned}\phi^f &= [ \sum_{k \in J} \phi^k ] / K_{f a i l} , \\ \phi^p &= [ \sum_{k \notin J} \phi^k ] / K_{p a s s} ,\end{aligned}\tag{5.13}$$

where  $J$  is the index set identifying the failed samples.  $K_{f a i l}$  and  $K_{p a s s}$  are the numbers of failed and passed samples, respectively. The nominal point  $\phi^0$  is then adjusted along the direction given by  $s = \phi^p - \phi^f$  using a line search. This algorithm is simple but also heuristic. The relationship between the gravity centers and the yield is not clear in a general multidimensional problem.

### 5.3.4 The Parametric Sampling Method

The parametric sampling method by Singhal and Pinel (1981) has provided another promising direction. A continuous estimate of yield (as opposed to the Monte Carlo estimate using discrete samples) is given by the following integral

$$Y(\mathbf{x}) = \int_{-\infty}^{+\infty} I_a(\boldsymbol{\phi}) \Gamma(\boldsymbol{\phi}, \mathbf{x}) d\boldsymbol{\phi} , \quad (5.14)$$

where  $I_a(\boldsymbol{\phi})$  is the acceptance index defined in (5.4) and  $\Gamma(\boldsymbol{\phi}, \mathbf{x})$  the parameter distribution density function which depends on the design variables  $\mathbf{x}$  (e.g., the nominal point specifies the mean value and the tolerances control the standard deviations). Normally, in order to estimate the yield, we generate samples  $\boldsymbol{\phi}^k$ ,  $k = 1, 2, \dots, K$ , from the component density  $\Gamma$ , perform  $K$  circuit analyses and then take the average of  $I_a(\boldsymbol{\phi}^k)$ . For each new set of variables  $\mathbf{x}$  we would have a new density function and, therefore, the sampling and circuit analyses have to be repeated.

The approach of parametric sampling is based on the concept of importance sampling as

$$Y(\mathbf{x}) = \int_{-\infty}^{+\infty} I_a(\boldsymbol{\phi}) \frac{\Gamma(\boldsymbol{\phi}, \mathbf{x})}{h(\boldsymbol{\phi})} h(\boldsymbol{\phi}) d\boldsymbol{\phi} , \quad (5.15)$$

where  $h(\boldsymbol{\phi})$  is called the sampling density function. The samples  $\boldsymbol{\phi}^k$  are generated from  $h(\boldsymbol{\phi})$  instead of  $\Gamma(\boldsymbol{\phi}, \mathbf{x})$ . An estimate of the yield is made as

$$Y(\mathbf{x}) \approx \frac{1}{K} \sum_{k=1}^K I_a(\boldsymbol{\phi}^k) \frac{\Gamma(\boldsymbol{\phi}^k, \mathbf{x})}{h(\boldsymbol{\phi}^k)} = \frac{1}{K} \sum_{k=1}^K I_a(\boldsymbol{\phi}^k) W(\boldsymbol{\phi}^k, \mathbf{x}) . \quad (5.16)$$

The weights  $W(\boldsymbol{\phi}^k, \mathbf{x})$  compensate for the use of a sampling density different from the component density.

This approach has two clear advantages. Firstly, once the acceptance indices  $I_a(\boldsymbol{\phi}^k)$  are calculated, no more model simulations are required when  $\mathbf{x}$  is changed. Furthermore, if  $\Gamma$  is a differentiable density function, then gradients of the estimated yield are

readily available. Hence, powerful optimization techniques may be employed. In practice the algorithm starts with a large number of base points sampled from  $h(\phi)$  to construct the initial databank. To maintain a sufficient accuracy, the databank needs to be updated by adding new samples during optimization.

This approach, however, is not applicable to non-differentiable density functions such as uniform, discrete and truncated distributions. It can be extended to include tunable parameters if the tuning ranges are fixed or practically unlimited. In this case the index  $I_a(\phi^k)$  is defined as 1 if  $\phi^k$  is acceptable after tuning. If  $\phi^k$  is not acceptable before tuning, then whether it can be tuned and, if so, by how much will have to be determined through optimization (which is separate from the optimization of yield). A variable tuning range (in order to minimize the cost) can not be accommodated.

#### 5.4 A GENERALIZED $\ell_p$ CENTERING ALGORITHM

In this section, we propose a generalized  $\ell_p$  centering algorithm which encompasses, in a unified formulation, problems of 100% yield (worst-case design) and less than 100% yield.

##### 5.4.1 Representing an Outcome by an $\ell_p$ Function

First we consider the centering problem where we have fixed tolerances and no tuning. The optimization variable vector  $\mathbf{x}$  contains elements of the nominal parameters  $\phi^0$ . Define

$$\mathbf{f} = [\mathbf{e}^T(\phi^1) \ \mathbf{e}^T(\phi^2) \ \dots \ \mathbf{e}^T(\phi^k)]^T \quad (5.17)$$

as the set of multi-circuit error functions. We can achieve a worst-case minimax design by

$$\underset{\mathbf{x}}{\text{minimize}} U(\mathbf{x}) = H_{\infty}(\mathbf{f}) = \max_k \max_j \{e_j(\phi^k)\}, \quad (5.18)$$

where the multiple circuits  $\phi^k$  are related to  $\phi^0$  by (5.3).

If a 100% yield is not possible, we would naturally look for a solution where the specifications are met by as many points (out of  $K$  circuits) as possible. For this purpose minimax is not a proper choice, since unless and until the worst case is dealt with nothing else seems to matter. We may attempt to substitute for  $H_{\infty}(\mathbf{f})$  a generalized  $\ell_2$  or  $\ell_1$  function, i.e.,  $H_2(\mathbf{f})$  or  $H_1(\mathbf{f})$ , hoping to reduce the emphasis given to the worst case. However, in (5.18) each outcome ( $\phi^k$ ) is represented by a set of discrete error functions and each error function makes a separate contribution to the  $\ell_p$  objective, thus obscuring the relationship between different outcomes.

The problem is better represented by finding, for each  $\phi^k$ , a scalar function which indicates directly whether  $\phi^k$  satisfies or violates the specifications and by how much. For this purpose, we choose a set of generalized  $\ell_p$  functions as

$$v_k(\mathbf{x}) = H_p(e(\phi^k)), \quad k = 1, 2, \dots, K. \quad (5.19)$$

The sign of  $v_k$  indicates the acceptability of  $\phi^k$  and the magnitude of  $v_k$  measures, so to speak, the distance between  $\phi^k$  and the boundary of the acceptable region. For example, when  $p = \infty$  it is a worst-case measure whereas when  $p = 2$  it becomes closer to the Euclidean norm.

### 5.4.2 Generalized $\ell_p$ Centering

We define a generalized  $\ell_p$  centering as

$$\underset{\mathbf{x}}{\text{minimize}} U(\mathbf{x}) = H_p(\mathbf{u}(\mathbf{x})), \quad (5.20)$$

where  $\mathbf{x}$  contains the nominal (center) point, and

$$\mathbf{u}(\mathbf{x}) = \begin{bmatrix} \alpha_1 v_1 \\ \vdots \\ \alpha_k v_k \end{bmatrix} = \begin{bmatrix} \alpha_1 H_q(\mathbf{e}(\phi^1)) \\ \vdots \\ \alpha_k H_q(\mathbf{e}(\phi^k)) \end{bmatrix} \quad (5.21)$$

and  $\alpha_1, \dots, \alpha_k$  are a set of positive multipliers. Different values of  $p$  and  $q$  lead to different variations of the algorithm. We shall discuss separately the case where a nonpositive  $U(\mathbf{x})$  exists and the case where we always have  $U(\mathbf{x}) > 0$ , noticing that for a given  $\mathbf{x}$  the sign of  $U(\mathbf{x})$  does not depend on  $p$ ,  $q$  or any  $\alpha_k$ .

In the first case, the existence of a  $U(\mathbf{x}) \leq 0$  indicates that a 100% yield is attainable. Since the sign of  $U(\mathbf{x})$  is independent of  $p$ ,  $q$  and  $\alpha$ , we should be able to achieve a  $U(\mathbf{x}) \leq 0$  (i.e., a 100% yield) using any values for  $p$ ,  $q$  and  $\alpha$ . However, the specific solution  $\mathbf{x}$  at which  $U(\mathbf{x})$  attains its minimum depends on  $p$ ,  $q$  and  $\alpha$ . It means that the centered design  $\phi^0$  will be influenced by the choice of  $p$ ,  $q$  and  $\alpha$ . Interestingly, the minimax worst-case centering becomes a special case by letting both  $p, q = \infty$  and using unit multipliers.

Now consider the case where the optimal yield is less than 100%. In this case the set  $J = \{k \mid u_k > 0\}$  is not empty and the generalized  $\ell_p$  objective is defined by the one-sided  $\ell_p$  function  $H_p^+$  (see (2.8) of Section 2.2.1). To emphasize the effect of  $p$ , we write the one-sided  $\ell_p$  sum as

$$U(\mathbf{x}, p) = \sum_{k \in J} [u_k(\mathbf{x})]^p . \quad (5.22)$$

As  $p$  decreases towards zero,  $U(\mathbf{x}, p)$  approaches the total number of unacceptable outcomes which is also given by  $K(1 - Y(\mathbf{x}))$ ,  $Y(\mathbf{x})$  being the discrete yield. We may, therefore, attempt to minimize  $U(\mathbf{x}, p)$  as a means of increasing  $Y(\mathbf{x})$ . However, for  $p < 1$   $U(\mathbf{x}, p)$  is not a convex function. The smallest  $p$  that preserves the convexity is 1, which leads to the one-sided  $\ell_1$  function as

$$H_1^+(\mathbf{u}) = \sum_{k \in J} u_k(\mathbf{x}) = \sum_{k \in J} \alpha_k v_k(\mathbf{x}) . \quad (5.23)$$

If we choose the multipliers as

$$\alpha_k = 1/v_k(\mathbf{x}_r) , \quad (5.24)$$

where  $\mathbf{x}_r$  is a reference point, then  $H_1^+$  defined in (5.23) becomes precisely  $K(1 - Y(\mathbf{x}))$  at  $\mathbf{x}_r$ . From this point of view, (5.23) and (5.24) define a smooth and convex interpolation function for the purpose of enhancing the discrete yield. The one-sided  $\ell_1$  algorithm developed in Chapter 2 is a powerful tool for minimizing  $H_1^+$ .

From a different viewpoint, we know that a stationary point  $\mathbf{x}^*$  of  $U(\mathbf{x}, p)$  as given by (5.22) is characterized by the derivatives of  $[u_k(\mathbf{x}^*)]^p$  which is

$$u_k^{p-1} \mathbf{u}'_k = (\alpha_k v_k)^{p-1} \alpha_k \mathbf{v}'_k = \alpha_k^p v_k^{p-1} \mathbf{v}'_k . \quad (5.25)$$

A stationary point of  $H_1^+$  as given by (5.23) is characterized by

$$\mathbf{u}'_k = \alpha_k \mathbf{v}'_k . \quad (5.26)$$

In order for  $\mathbf{x}^*$  to be a stationary point of  $H_1^+$  also, we equate (5.25) to (5.26), resulting in the multipliers given by (5.24) with the reference point  $\mathbf{x}_r = \mathbf{x}^*$ . In other words, by using appropriate multipliers, we may be able to approximate the minimizer of  $U(\mathbf{x}, p)$

by solving a convex  $\ell_1$  problem. Minimizing  $U(\mathbf{x}, \mathbf{p})$  is in turn an approximation to maximizing  $Y(\mathbf{x})$ .

In practice, we can not know  $\mathbf{x}^*$  prior to the optimization. We may take the starting point  $\mathbf{x}_0$  as a convenient reference point and, consequently, define the multipliers by

$$\alpha_k = \begin{cases} 1/v_k(\mathbf{x}_0), & \text{if } v_k(\mathbf{x}_0) > \sigma, \\ 1/\sigma, & \text{otherwise,} \end{cases} \quad k = 1, 2, \dots, K, \quad (5.27)$$

where  $\sigma$  is a small positive number suitably chosen for a numerical implementation. Notice that by definition the magnitude of  $v_k$  measures the closeness of  $\phi^k$  to the acceptable region. A small  $v_k$  indicates that  $\phi^k$  is close to be satisfying the specifications. Therefore, we assign a large multiplier to it so that more emphasis will be given to  $\phi^k$  during optimization. On the other hand, we deemphasize those points that are far away from the boundary of the acceptable region because their contributions to the yield are less likely to change.

The roles of the multipliers can also be discussed in the light of the penalty function approach. If we consider  $\mathbf{v} \leq \mathbf{0}$  as a set of nonlinear constraints, then the maximizer of the discrete yield will have a maximum number of constraints satisfied. The generalized  $\ell_p$  centering is like a penalty function method. If a 100% yield is possible, then the centered solution will satisfy all the constraints, regardless of the values of multipliers. Otherwise, the yield is less than 100% and the optimum satisfies only a subset of the constraints. In this case if we assign sufficiently large multipliers to the constraints in the optimal subset, then

the  $\ell_1$  objective (5.23) becomes an exact penalty function. In the present algorithm, we consider those outcomes that are close to the acceptable region as likely candidates for the optimal subset. We also consider  $1/\sigma$ , in (5.27), as a sufficient large bound for the multipliers.

A sequential process may also be constructed in which we solve (5.20), update the multipliers at the solution and repeat the optimization. Bearing in mind that design centering should always be initiated from a good nominal solution and, therefore, drastic changes in the variables are unlikely to take place, we can expect such a sequential process to be a stable one. Incidentally, the generalized  $\ell_p$  centering algorithm can certainly be applied to a single circuit to obtain a nominal design. In that sense, the new algorithm is a natural extension to the  $\ell_p$  nominal circuit optimization. If we can not achieve even an acceptable nominal point, then any attempt to optimize the yield is doomed to failure.

#### 5.4.3 Tolerancing and Tuning

The generalized  $\ell_p$  centering algorithm can be extended to accommodate considerations of tolerancing and tuning.

We need to define a function or functions which appropriately relate the tolerances and tuning to the cost of production. One possible choice of such a cost function is, similar to (5.11),

$$C(\mathbf{x}) = \sum_{i \in I_\epsilon} \frac{a_i}{\epsilon_i} + \sum_{i \in I_t} b_i t_i - C_0, \quad (5.28)$$

where the index sets  $I_\epsilon$  and  $I_t$  identify the toleranced and tunable

parameters, respectively.  $a_i$  and  $b_i$  are nonnegative weights.  $C_0$  is a realistic target for the cost. It is also possible to define several cost functions representing separate trade-offs in a complex environment such as VLSI design (similar to the multiple objectives described by Brayton, Hachtel and Sangiovanni-Vincentelli (1981)). For instance, (5.28) may be broken up into multiple cost functions associated with separate groups of parameters.

By this formulation, the cost function is treated in the same way as we treat  $\mathbf{e}(\boldsymbol{\phi})$ . When  $C(\mathbf{x}) \leq 0$ , we say that  $\mathbf{x}$  satisfies the specification ( $C_0$ ) on the cost. In fact, we can adjust  $C_0$  and the weights  $a_i$ ,  $b_i$  so that  $C(\mathbf{x})$  is made comparable in value with the error functions.

The optimization problem is defined as

$$\underset{\mathbf{x}}{\text{minimize}} \quad U(\mathbf{x}) = H_p(\mathbf{u}(\mathbf{x})), \quad (5.29)$$

where the variables  $\mathbf{x}$  include the nominal point  $\boldsymbol{\phi}^0$ , the tolerances  $\epsilon_i$ ,  $i \in I_\epsilon$ , and the tuning adjustments  $\tau_i^k$ ,  $|\tau_i^k| \leq t_i$ ,  $k = 1, 2, \dots, K$ ,  $i \in I_t$ , for the multiple circuits defined by (5.9). We also have

$$\mathbf{u}(\mathbf{x}) = \begin{bmatrix} \alpha_1 v_1 \\ \vdots \\ \alpha_k v_k \\ \alpha_{k+1} v_{k+1} \end{bmatrix} = \begin{bmatrix} \alpha_1 H_q(\mathbf{e}(\boldsymbol{\phi}^1)) \\ \vdots \\ \alpha_k H_q(\mathbf{e}(\boldsymbol{\phi}^k)) \\ \alpha_{k+1} C(\mathbf{x}) \end{bmatrix}. \quad (5.30)$$

Following the discussions of the last section, we can solve (5.30) for a solution, say  $\mathbf{x}^*$ . If  $U(\mathbf{x}^*) \leq 0$  then the design specifications have all been satisfied and the target for cost has been met. If, on the other hand,  $U(\mathbf{x}^*) > 0$ , we can conclude that either the design specifications are too tight or the target cost is un-

realistic or perhaps both.

#### 5.4.4 Circuit Examples

In this section, we present two circuit examples to illustrate the usefulness of the theory.

The first example was originally considered by Singhal and Pinel (1981). The circuit involved is a Chebyshev lowpass filter as shown in Fig. 5.6. Fifty-one frequencies {0.02, 0.04, ..., 1.0, 1.3 Hz} are considered. An upper specification of 0.32dB on the insertion loss is defined from 0.02 to 1.0Hz, and a lower specification of 52dB on the insertion loss is defined at 1.3Hz.

Valuable results on the statistical design of this circuit have been reported by Singhal and Pinel (1981) using the parametric sampling method. Normally distributed tolerances were assumed for the 11 circuit components. Starting from a nominal point obtained by standard filter synthesis, a worst-case design as well as a minimum cost design were achieved.

As we have pointed out in Section 5.3.4, the parametric sampling method is not applicable to non-differentiable (such as the uniform) distributions. Here, we apply the generalized  $\ell_p$  (with  $p = 1$ ) centering algorithm to the Chebyshev lowpass filter, assuming a uniformly distributed 1.5% relative tolerance for each component. The nominal design by standard synthesis was used as a starting point. It has a yield of 49%. A yield of 84% is achieved by our centered solution which involved a sequence of three design cycles with a total CPU time of 66 seconds on the VAX 8600. The

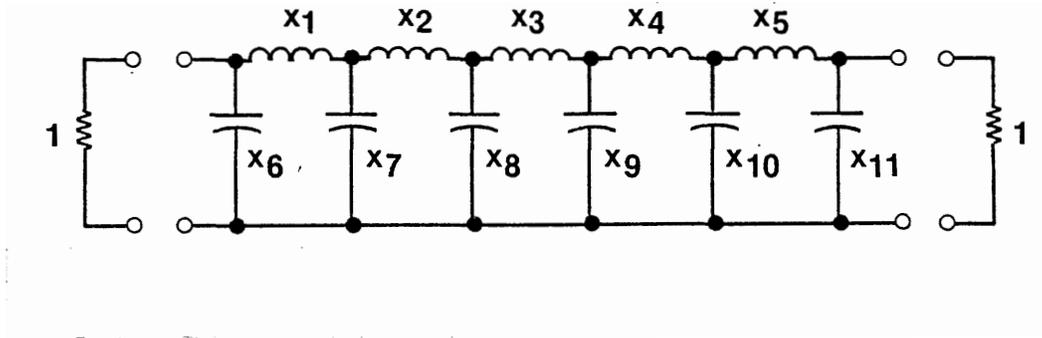


Fig. 5.6 The Chebyshev lowpass filter (Singhal and Pintel 1981).

one-sided  $l_1$  algorithm developed in Chapter 2 was employed to solve the optimization problem. Some details are provided in Table 5.1.

For the second example, we consider the design centering of multi-coupled cavity filters. The nominal optimization of these filters has been described in detail in Section 4.3. This example involves a 6th-order filter. The center frequency is 4GHz and the bandwidth 40MHz. We have for the passband (3980MHz - 4020MHz) a lower specification of 20dB on the return loss and for the stopband (below 3976MHz or above 4024MHz) a lower specification of 20dB on the insertion loss. Five couplings of the filter as well as the input and output transformer ratios are considered as variables.

With respect to the given specifications, a minimax nominal design was obtained at which the yield is 41%, assuming a uniformly distributed 2.5% relative tolerance for each variable. Using our centering algorithm a 65% yield was achieved, after two design cycles and a total CPU time of 23 seconds on the VAX 8600. The details are given in Table 5.2.

In (5.27), for numerical reasons, we have defined a bound on the multipliers. In the examples of this section,  $\sigma = 0.01$  was used (i.e., the bound on the multipliers was 100).

We conclude this section with a discussion on the computational efficiency related to our centering algorithm. Notice that as long as the yield is less than 100%, negative error functions do not contribute to the generalized  $l_p$  objective. Therefore, under certain conditions, the circuit analyses corresponding to these error functions can be saved. Such a saving becomes increasingly

TABLE 5.1

GENERALIZED  $\ell_1$  CENTERING OF THE CHEBYSHEV LOWPASS FILTER

Component	Nominal Values			
	Case 1	Case 2	Case 3	Case 4
$x_1$	0.2251	0.21954	0.21705	0.21530
$x_2$	0.2494	0.25157	0.24677	0.23838
$x_3$	0.2523	0.25529	0.24784	0.24120
$x_4$	0.2494	0.24807	0.24019	0.23687
$x_5$	0.2251	0.22042	0.21753	0.21335
$x_6$	0.2149	0.22628	0.23565	0.23093
$x_7$	0.3636	0.36739	0.37212	0.38224
$x_8$	0.3761	0.36929	0.38012	0.39023
$x_9$	0.3761	0.37341	0.38370	0.39378
$x_{10}$	0.3636	0.36732	0.37716	0.38248
$x_{11}$	0.2149	0.22575	0.22127	0.23129
Yield	49%	78%	80%	84%
Number of circuits		50	100	100
Starting point		Case 1	Case 2	Case 3
Number of iterations		16	18	13
CPU time (VAX 8600)		10 sec.	30 sec.	26 sec.

A uniformly distributed 1.5% relative tolerance is assumed for each component. The yield in this table was estimated by Monte Carlo analyses with 300 samples. The parameter values in Case 1 were obtained by standard filter synthesis (Singhal and Pinel 1981).

TABLE 5.2

GENERALIZED  $\ell_1$  CENTERING OF A MULTI-COUPLED CAVITY FILTER

Component	Nominal Values		
	Case 1	Case 2	Case 3
$M_{12} = M_{21} = M_{56} = M_{65}$	0.93240	0.93795	0.93728
$M_{23} = M_{32} = M_{45} = M_{54}$	0.52648	0.51946	0.50747
$M_{34} = M_{43}$	0.88160	0.90228	0.92950
$M_{16} = M_{61}$	0.16173	0.16798	0.26720
$M_{25} = M_{52}$	-0.44036	-0.46782	-0.54209
$n_1 = n_2$	1.25824	1.25778	1.27982
Yield	41%	55%	65%
Number of circuits		100	100
Starting point		Case 1	Case 2
Number of iterations		20	17
CPU time (VAX 8600)		11 sec.	12 sec.

A uniformly distributed 2.5% relative tolerance is assumed for each component. The yield in this table was estimated by Monte Carlo analyses with 300 samples. The parameter values in Case 1 were obtained by a minimax nominal optimization.

significant as the yield is improved (since more functions become negative). This technique has been incorporated in our implementation. More specifically, at the  $i$ th iteration, we calculate a first-order estimate by

$$\bar{e}_j(\phi_1^k) = e_j(\phi_{1-1}^k) + \mathbf{g}^T(\phi_1^k - \phi_{1-1}^k), \quad (5.31)$$

where  $\phi_1^k$  denotes the parameters for the  $k$ th circuit at the  $i$ th iteration and  $\mathbf{g}$  contains the sensitivities of  $e_j$  with respect to the circuit parameters. If this estimate  $\bar{e}_j < -\delta$ , where  $\delta$  is a small positive number, then the computation of  $e_j(\phi_1^k)$  is saved and  $\bar{e}_j$  is used as an approximation, otherwise an exact simulation is required. The sensitivities  $\mathbf{g}$  are updated whenever a relevant exact simulation is performed or otherwise kept constant.  $\delta$  is intended as a safe margin to allow for the accumulated errors in the first-order estimation. A suitable value may depend on each application. In the examples we have set it to 10% of the corresponding specification. The computational saving realized by this method can be significant. Without the benefit of this technique, it took 203 seconds, instead of 66 seconds, to achieve the same result for the first example.

## 5.5 CONCLUDING REMARKS

Realistic approaches to circuit design have been studied in this chapter. The common aim of all these approaches has been to improve the yield and reduce the production cost in the presence of tolerances and uncertainties. The various possible uncertainties associated with circuit models at different levels have been iden-

tified. The concept of multiple circuits has been defined and related to yield estimation, design centering, optimal tolerancing and tuning. We have also reviewed some important techniques in statistical design including the worst-case approach, the simplicial and multidimensional approximations, the gravity method and the parametric sampling method.

A generalized  $\ell_p$  centering algorithm has been proposed as a natural extension to the  $\ell_p$  nominal circuit optimization and the minimax worst-case design. It has provided a unified approach to design centering and yield enhancement. The theoretical implications of the proposed formulation have been discussed. Also, two circuit examples have illustrated the practical usefulness of the algorithm. The incorporation of suitable simulation saving techniques, such as the one that we have introduced, as well as the rapid progress in mass computation hardware will certainly further reduce the expense of statistical design.

## CHAPTER 6

### A MULTI-CIRCUIT APPROACH TO DEVICE MODELING

#### 6.1 INTRODUCTION

Device modeling essentially involves approximating measured responses of a system by calculated responses using a suitable model. It is also related to verifying a proposed model based on its consistency with respect to physical perturbations. Also, we may employ modeling techniques to establish an analytical relationship between the physical parameters and the parameters of an equivalent circuit.

The traditional approach to modeling is almost entirely directed at achieving the best possible match between the measured and calculated responses. When the presence of uncertainties causes an imperfect match between model responses and measurements or a family of nonunique solutions (with respect to the responses selected), the traditional approach has serious shortcomings.

Recently, a novel approach to robust device modeling was presented by Bandler, Chen and Daijavad (1986b). It exploits the unique properties of the  $l_1$  optimization and employs the concept of simultaneous processing of multiple circuits. It has the advantage of establishing not only a good circuit model whose responses match as much as possible the measurement, but also a reliable measure of the self-consistency of the model.

This chapter is devoted to describing the motivation, the

theoretical foundation and the practical applications of this new approach. The unique properties of the  $\ell_1$  optimization are discussed and a superlinearly convergent  $\ell_1$  algorithm described by Bandler, Kellermann and Madsen (1987) is employed. We show that the use of multiple circuits may increase the identifiability of a problem, leading to the identification of a unique set of model parameters. An illustrative simple RC circuit as well as an actual FET device are provided as examples.

The new approach is also applied to automatic model verification, where the consistency of the identified model parameters with respect to physical perturbations is checked. If successful, the method provides confidence in the proposed model; otherwise it proves the model's incorrectness. This technique is demonstrated by an 8th-order multi-coupled cavity filter example.

The multi-circuit formulation is also employed to establish a local relationship between the physical parameters and the equivalent circuit parameters for a 6th-order cavity filter. Such an experimental relationship will be extremely useful in guiding post-production tuning of an actual device.

## 6.2 EXTENSION OF THE NOMINAL DESIGN CONCEPT TO MODELING

Traditionally, the approach of nominal design has been extended to solving modeling problems. A set of measurements made on the physical system, denoted by  $\mathbf{F}^M$ , serves as single specifications. Error functions are created from the differences between the calculated responses  $\mathbf{F}(\boldsymbol{\phi}^0)$  and the measured responses  $\mathbf{F}^M$ ,

similarly to (4.3), as

$$e_j(\boldsymbol{\phi}^0) = w_j |F_j(\boldsymbol{\phi}^0) - F_j^M|, \quad j = 1, 2, \dots, m. \quad (6.1)$$

By minimizing an  $\ell_p$  norm of the error functions, we attempt to identify a set of model parameters  $\boldsymbol{\phi}^0$  such that  $F(\boldsymbol{\phi}^0)$  best matches  $F^M$ .

Such a casual treatment of modeling as if it were a special case of design is often unjustifiable, due to the lack of consideration to the uniqueness of the solution. In circuit design, one satisfactory nominal point, possibly out of many feasible solutions (i.e., not unique), will suffice. In modeling, however, the uniqueness of the solution is almost always an essential part of the problem. Affected by the uncertainties at many levels as well as unavoidable measurement errors, the model obtained by a nominal optimization is often nonunique and unreliable.

### 6.3 MODELING USING MULTIPLE CIRCUITS

#### 6.3.1 Uncertainties that Affect Modeling

For the convenience of discussion, we replicate the notation introduced in Section 5.2.1 for tolerances and uncertainties as

$$\begin{aligned} \mathbf{F}^M &= \mathbf{F}^{M,0}(\boldsymbol{\phi}^M) + \Delta\mathbf{F}^M, \\ \boldsymbol{\phi}^M &= \boldsymbol{\phi}^{M,0} + \Delta\boldsymbol{\phi}^M, \\ \mathbf{F}^H &= \mathbf{F}^{H,0}(\boldsymbol{\phi}^L) + \Delta\mathbf{F}^H, \\ \mathbf{F}^L &= \mathbf{F}^{L,0}(\boldsymbol{\phi}^H) + \Delta\mathbf{F}^L, \\ \boldsymbol{\phi}^H &= \boldsymbol{\phi}^{H,0}(\boldsymbol{\phi}^L) + \Delta\boldsymbol{\phi}^H, \\ \boldsymbol{\phi}^L &= \boldsymbol{\phi}^{L,0} + \Delta\boldsymbol{\phi}^L, \end{aligned} \quad (6.2)$$

where the superscript 0 identifies a nominal value and the prefix  $\Delta$  indicates an associated uncertainty or inaccuracy.

The adverse effect of these uncertainties on modeling can be discussed under the following categories.

1) Measurement errors will inevitably exist in practice, as represented by  $\Delta F^M$  in (6.2).

2) Even without measurement errors, the calculated response  $F^{H,0}$  may never be able to match  $F^{M,0}$  perfectly, due to, e.g., the use of a simulation model of insufficient order or inadequate complexity. Such an inherent mismatch is accounted for by  $\Delta F^H$ .

3) Even if neither  $\Delta F^M$  nor  $\Delta F^H$  exists so that  $F^{H,0} = F^M$ , we may still not be able to uniquely identify the model parameters from the set of measurements that has been selected. This happens when the system of (generally nonlinear) equations  $F^{H,0}(\phi) - F^M = 0$  is underdetermined. It typically occurs when, for any reason, many internal nodes are inaccessible to direct measurement. An over-complicated equivalent circuit model is often at the heart of this phenomenon.

4) The parasitic effects that are not adequately modeled by  $\phi^H$  contribute to the uncertainty  $\Delta F^L$ , which becomes another source of interference with the modeling process.

Consider the case in which modeling is applied to obtain a suitable  $\phi$  such that  $F^H(\phi)$  approximates  $F^M$ . The nominal circuit approach, as described in Section 6.2, may be able to cope with measurement errors and the uncertainties due to the use of a lower-order model, identified by 1) and 2), and comes up with a  $\phi$  which

minimizes  $\Delta F^M$  and  $\Delta F^H$  in a certain sense. But it will not be able to overcome the problem of nonunique solutions.

Often in practice, we are not able to determine unambiguously the identifiability of a system and the uniqueness of a model, because all the uncertainties described above can be present at the same time. There will be, typically, a family of solutions which all produce reasonable and similar matches between the measured and the calculated responses. We can not, therefore, rely on any particular set of parameters.

### 6.3.2 Multiple Circuits and Common Variables

The use of multiple circuits in device modeling was originally considered by Bandler, Chen and Daijavad (1986b). Multiple circuits are created by making deliberate adjustments on some of the physical parameters. For example, we can change the biasing conditions of an active device and obtain multiple sets of measurements. By doing so, we introduce perturbations to the model which cause some parameters in  $\phi$  to change by an unknown amount. For this approach to be successful, each physical adjustment should produce changes in only a few parameters in  $\phi$ .

Although we do not know the changes in  $\phi$  quantitatively, it is often possible to identify which model parameters may have been affected by each of the physical adjustments. Such a qualitative knowledge may be apparent from the definition of the model or may come from practical experience. In the attempt to process the multiple circuits simultaneously, we define those model parameters

that are not supposed to change as common variables and, at the same time, allow the others to vary between different circuits. By doing so, we subject the solution to the anticipated topological constraints. In other words, from a family of possible solutions we give preference to the one that exhibits the desired consistency with respect to physical perturbations, thus increasing the reliability of the result.

To formulate this mathematically, let

$$\phi^k = \begin{bmatrix} \phi_c^k \\ \phi_a^k \end{bmatrix}, \quad (6.3)$$

where  $\phi_c^k$  contains the common variables and  $\phi_a^k$  contains the independent variables, i.e., model parameters that are allowed to vary between the  $k$ th circuit and the reference circuit  $\phi^0$ . We then define the optimization variables by

$$\mathbf{x} = \begin{bmatrix} \phi^0 \\ \phi_a^1 \\ \cdot \\ \cdot \\ \phi_a^K \end{bmatrix} \quad (6.4)$$

and state the optimization problem as to

$$\underset{\mathbf{x}}{\text{minimize}} \quad U(\mathbf{x}) = \|\mathbf{f}\|_p, \quad (6.5)$$

where

$$\mathbf{f} = [\mathbf{e}^T(\phi^0) \quad \mathbf{e}^T(\phi^1) \quad \dots \quad \mathbf{e}^T(\phi^K)]^T. \quad (6.6)$$

The concept of common and independent variables is depicted in Fig. 6.1.

Now, suppose that we do not have a clear idea as to which model parameters may have been affected by the adjustment on  $\phi^M$ .

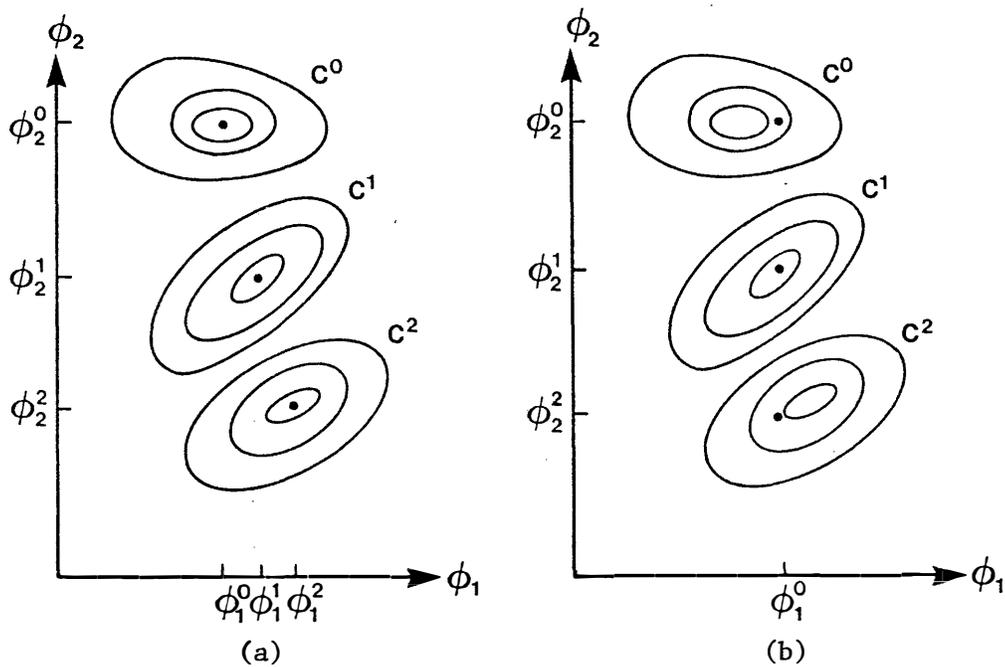


Fig. 6.1 An illustration of common and independent variables in multi-circuit modeling. Three circuits are created by making two physical adjustments. Assume that we know that  $\phi_1$  should not be affected by the physical adjustments.  $C^0$ ,  $C^1$  and  $C^2$  are contours of the error functions corresponding to the three circuits.

(a) The three circuits are processed separately.  $\phi_1^0$ ,  $\phi_1^1$  and  $\phi_1^2$  turn out to have different values (which is inconsistent with our knowledge) due to uncertainties.

(b) Consistent results are obtained by defining  $\phi_1$  as a common variable and processing three circuits simultaneously.

In this case, we let

$$\mathbf{x} = \begin{bmatrix} \phi^0 \\ \phi^1 \\ \cdot \\ \cdot \\ \phi^K \end{bmatrix} \quad (6.7)$$

and change the objective function to an  $\ell_p$  norm of

$$\mathbf{f} = \begin{bmatrix} e(\phi^0) \\ \cdot \\ \cdot \\ e(\phi^K) \\ \alpha_1(\phi^1 - \phi^0) \\ \cdot \\ \cdot \\ \alpha_K(\phi^K - \phi^0) \end{bmatrix}, \quad (6.8)$$

where  $\alpha_1, \alpha_2, \dots, \alpha_K$  are nonnegative multipliers (weights).

Using this formulation, while minimizing the errors  $e$ , we penalize the objective function for any deviates between  $\phi^k$  and  $\phi^0$ , since our only available knowledge is that only a few parameters in  $\phi^k$  should have any significant changes.

### 6.3.3 Computational Considerations

A brief discussion on computational considerations, related to coordinating the multi-circuit formulation with an optimizer, is in order. In the process of solving the optimization problem given by (6.5), a new set of values is obtained for  $\mathbf{x}$  at each iteration. The following steps should then take place.

Step 1 Recover from  $\mathbf{x}$  the multiple sets of circuit parameters, namely  $\phi^k$ ,  $k = 1, 2, \dots, K$ .

Comment If common variables are defined as in (6.4), then (6.3) should be used to construct  $\phi^k$ . Otherwise,  $\phi^k$  is simply a subset of  $\mathbf{x}$ , as in (6.7).

Step 2 Perform K circuit simulations to obtain the model responses  $F(\phi^k)$  as well as their sensitivities  $\partial F/\partial \phi^k$ .

Step 3 From the calculated and measured responses formulate the error functions  $e(\phi^k)$ , and from the sensitivities of F formulate the Jacobian matrices  $J^k = [\partial e^T(\phi^k)/\partial \phi^k]^T$ .

Step 4 Construct the multi-circuit error functions  $f$  according to (6.6) or (6.8) as appropriate. The Jacobian of  $f$  with respect to  $\mathbf{x}$  is given by

$$\left[\frac{\partial f^T}{\partial \mathbf{x}}\right]^T = \begin{bmatrix} J^0 A^0 \\ \cdot \\ \cdot \\ J^K A^K \end{bmatrix}$$

if (6.6) is used, where  $A^k = [\partial(\phi^k)^T/\partial \mathbf{x}]^T$ ,  $k = 1, \dots, K$ , are constant index matrices. If (6.8) is used, then

$$\left[\frac{\partial f^T}{\partial \mathbf{x}}\right]^T = \begin{bmatrix} J^0 & 0 & 0 & \dots & 0 \\ & & \dots & & \\ 0 & 0 & 0 & \dots & J^K \\ -\alpha_1 \mathbf{1} & \alpha_1 \mathbf{1} & 0 & \dots & 0 \\ & & \dots & & \\ -\alpha_K \mathbf{1} & 0 & 0 & \dots & \alpha_K \mathbf{1} \end{bmatrix},$$

where the multipliers  $\alpha_k$  are defined in (6.8).

Usually, the values of  $f$  and its Jacobian obtained after these steps are sufficient for the optimizer to proceed with the next iteration. Here, we have assumed that exact network sensitivities are made available using, e.g., the approach described in

Chapter 3. Efficient methods of gradient approximations will be presented in Chapter 7.

#### 6.3.4 Exploiting the Unique Property of the $\ell_1$ Optimization

The use of the  $\ell_1$  norm as compared to the other norms has a distinct property that some large components of  $\mathbf{f}$  are deemphasized, i.e., at the solution it allows for a few  $f_j$ 's which are much larger than the others. This means that a few gross measurement errors are better tolerated by the  $\ell_1$  than any other norm, as has been discussed by Bartels and Conn (1981). Furthermore, in the context of this chapter, the  $\ell_1$  norm is particularly suitable for the formulation given by (6.8). An  $\ell_1$  solution is most likely to produce fewer deviates between  $\phi^k$  and  $\phi^0$ , which is consistent with our assumption that only a few model parameters should change due to a deliberate adjustment on the physical device.

To illustrate the above assertion, consider the example used by Bandler, Chen and Daijavad (1986b). We wish to find the rational approximant of the form

$$K(\mathbf{x}) = \frac{x_1 + x_2\omega + x_3\omega^2}{1 + x_4\omega + x_5\omega^2} \quad (6.9)$$

to the function  $\sqrt{\omega}$  in the interval  $\omega \in [0,1]$  using 51 uniformly spaced sample points. Without errors in the data, both the  $\ell_1$  and  $\ell_2$  solutions to this problem give a virtually perfect match to the actual function ( $\sqrt{\omega}$ ) over the given interval. Let this be case A. We now introduce a few large deviations in the actual function in two separate cases. In case B,  $\sqrt{\omega}$  is replaced by zero at five

points, namely  $\omega = 0.2, 0.4, \dots, 1.0$ . In case C we use zero at 0.4 and 0.8, and one at 0.2 and 0.6. The  $\ell_1$  solutions in these two cases are consistent with the  $\ell_1$  solution in case A. On the other hand, the presence of large deviations has severely affected the  $\ell_2$  solutions, as clearly shown in Fig. 6.2.

This unique property of the  $\ell_1$  optimization has also found applications in fault-isolation in analog circuits (Bandler and Salama 1985a) as well as the functional approach to postproduction tuning (Bandler and Salama 1985b).

## 6.4 PRACTICAL APPLICATIONS

### 6.4.1 Unique Identification of Model Parameters

The concept of parameter identifiability of a system has been discussed by Bandler and Salama (1985a) in the context of fault diagnosis in analog circuits. We can apply this concept to device modeling.

A measure of identifiability of  $\phi$  from the system

$$\mathbf{e}(\phi) = \mathbf{F}(\phi) - \mathbf{F}^M = 0 \quad (6.10)$$

is determined by testing the rank of the Jacobian matrix

$$\mathbf{J} = [\partial \mathbf{e}^T / \partial \phi]^T = [\partial \mathbf{F}^T / \partial \phi]^T. \quad (6.11)$$

If  $\mathbf{J}$  is not of full rank, then  $\phi$  is not uniquely identifiable from the set of measurements that has been selected. Usually  $\mathbf{J}$  has more rows than columns (i.e., there are more functions than variables), therefore the rank of  $\mathbf{J}$  is equal to its column rank.

To emphasize the partitioning of the parameters into common and independent variables, as in (6.3), we define

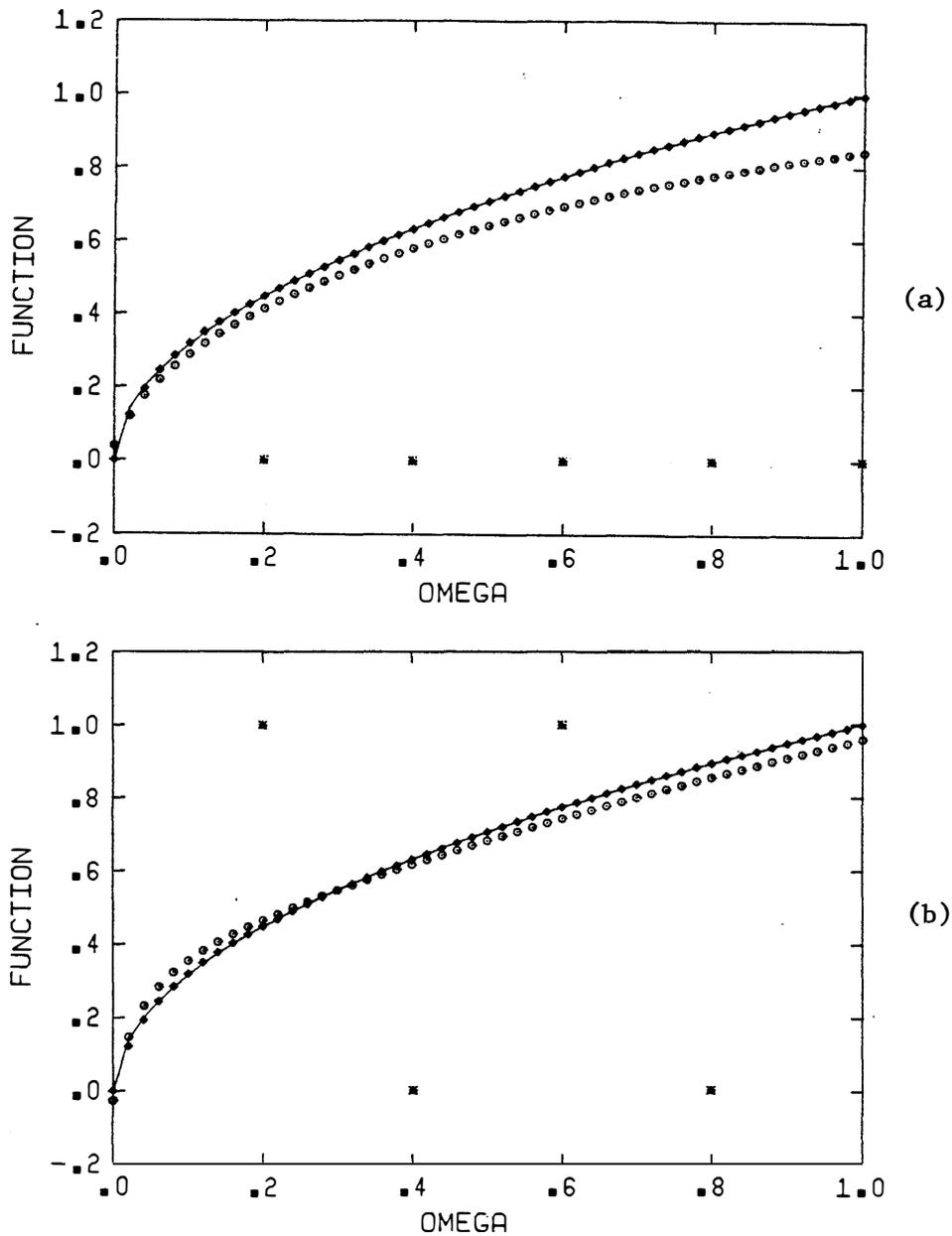


Fig. 6.2 Approximations using  $l_1$  and  $l_2$  optimization. The solid line is the actual function. Diamonds identify the approximation using  $l_1$  and circles represent approximations with  $l_2$ . Stars represent data points after large deliberate deviations. (a) and (b) correspond, respectively, to cases B and C.

$$\begin{aligned} \mathbf{J}_c(\boldsymbol{\phi}) &= \mathbf{J}_c(\boldsymbol{\phi}_c, \boldsymbol{\phi}_a) = [\partial \mathbf{F}^T / \partial \boldsymbol{\phi}_c]^T, \\ \mathbf{J}_a(\boldsymbol{\phi}) &= \mathbf{J}_a(\boldsymbol{\phi}_c, \boldsymbol{\phi}_a) = [\partial \mathbf{F}^T / \partial \boldsymbol{\phi}_a]^T. \end{aligned} \quad (6.12)$$

Then  $\mathbf{J} = [\mathbf{J}_c \ \mathbf{J}_a]$ .

The Jacobian matrix corresponding to the multi-circuit formulation, namely  $\mathbf{G} = [\partial \mathbf{f}^T / \partial \mathbf{x}]^T$  with  $\mathbf{f}$  and  $\mathbf{x}$  given by (6.4) and (6.6), can be constructed as

$$\mathbf{G} = \begin{bmatrix} \mathbf{J}_c^0 & \mathbf{J}_a^0 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{J}_c^1 & \mathbf{0} & \mathbf{J}_a^1 & \dots & \mathbf{0} \\ & & \dots & & \\ \mathbf{J}_c^K & \mathbf{0} & \mathbf{0} & \dots & \mathbf{J}_a^K \end{bmatrix}, \quad (6.13)$$

where the superscript identifies different circuits.

Assuming that a unique solution is not possible by using a single circuit, i.e., the matrix  $\mathbf{J}$  is rank-deficient, we show two necessary conditions for  $\mathbf{G}$  to be of full rank. If these conditions are met, we may be able to increase the identifiability of the system under consideration by using multiple circuits.

#### Necessary Condition 1

The matrices  $\mathbf{J}_a^k = \mathbf{J}_a(\boldsymbol{\phi}_c, \boldsymbol{\phi}_a^k)$  must have full rank for  $k = 0, 1, \dots, K$ . Otherwise, suppose that for at least one  $k$   $\mathbf{J}_a^k$  is rank-deficient, then the corresponding submatrix of  $\mathbf{G}$ , namely

$$\begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{J}_a^k \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}$$

is rank-deficient and so is the matrix  $\mathbf{G}$ .

This necessary condition imposes a restriction on the selection of independent variables. It requires that each single set of measurements must be at least sufficient to identify  $\phi_a^k$  (with  $\phi_c^k$  fixed).

Necessary Condition 2

The matrix  $J_c(\phi_c, \phi_a)$  must be a function of  $\phi_a$ . Otherwise, the matrices  $J_c^0, J_c^1, \dots, J_c^K$  would be identical and, consequently, the following submatrix of G

$$\begin{bmatrix} J_c^0 & J_a^0 \\ J_c^1 & 0 \\ & \vdots \\ J_c^K & 0 \end{bmatrix}$$

would have the same rank as  $[J_c^0 \ J_a^0]$  which is the Jacobian for a single circuit. In that case, G would be rank-deficient if J is.

The second necessary condition states mathematically the fact that if the designated common and independent variables are completely decoupled with respect to the measurements, then we can not expect to increase the identifiability by using multiple circuits. This is a mild condition since in most cases the network responses are nonlinear functions in circuit parameters.

We now illustrate the theory by a simple RC circuit as shown in Fig. 6.3 (Bandler, Chen and Daijavad 1986b, also Daijavad 1986). The parameters to be identified are  $\phi = [R_1 \ C \ R_2]^T$ . If we have measurements only on

$$V_2 = \frac{sCR_1R_2}{1 + sC(R_1 + R_2)}, \quad (6.14)$$

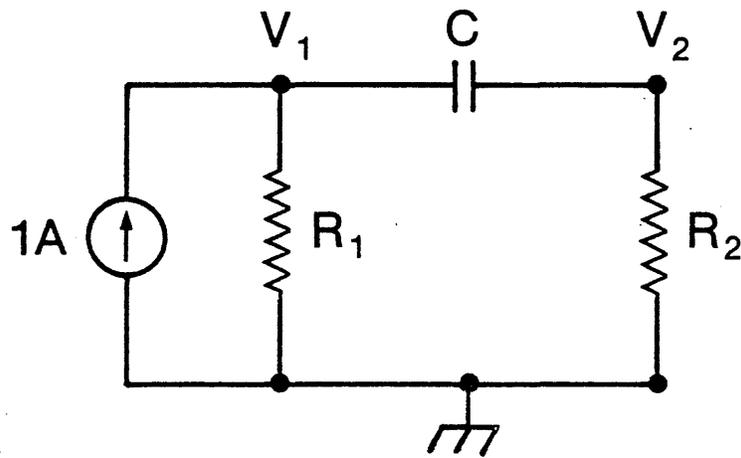


Fig. 6.3 Simple RC network.

where  $s = j\omega$ , it is clear by inspection that  $\phi$  can not be uniquely determined regardless of the number or choice of frequency samples, simply because  $R_1$  and  $R_2$  are observed in exactly the same way by  $V_2$ . Formally, we define the Jacobian matrix as

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \mathbf{F}}{\partial R_1} & \frac{\partial \mathbf{F}}{\partial C} & \frac{\partial \mathbf{F}}{\partial R_2} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{s_1 CR_2 (1+s_1 CR_2)}{[1+s_1 C(R_1+R_2)]^2} & \frac{s_1 R_1 R_2}{[1+s_1 C(R_1+R_2)]^2} & \frac{s_1 CR_1 (1+s_1 CR_1)}{[1+s_1 C(R_1+R_2)]^2} \\ \vdots & \vdots & \vdots \\ \frac{s_m CR_2 (1+s_m CR_2)}{[1+s_m C(R_1+R_2)]^2} & \frac{s_m R_1 R_2}{[1+s_m C(R_1+R_2)]^2} & \frac{s_m CR_1 (1+s_m CR_1)}{[1+s_m C(R_1+R_2)]^2} \end{bmatrix} \quad (6.15)$$

where  $\mathbf{F}$  contains  $V_2$  evaluated at  $m$  frequencies.

We find that  $\mathbf{J}$  can not have a rank greater than 2, since

$$\frac{\partial \mathbf{F}}{\partial R_1} + \frac{C(R_2 - R_1)}{R_1^2} \frac{\partial \mathbf{F}}{\partial C} - \frac{R_2^2}{R_1^2} \frac{\partial \mathbf{F}}{\partial R_2} = \mathbf{0}. \quad (6.16)$$

Therefore, a unique identification is impossible using only  $V_2$ .

Adjusting  $R_2$  by an unknown amount, we create a two-circuit modeling problem in which the variables, i.e., the parameters to be identified, are  $\mathbf{x} = [R_1^0 \ C^0 \ R_2^0 \ R_2^1]^T$ .

We proceed to show that in this example the necessary conditions for a unique parameter identification are satisfied.

The first necessary condition requires that for each  $k$  ( $k = 0$  and  $1$  in the present example)  $\mathbf{J}_a^k$  has full rank. Since

$$\mathbf{J}_a^k = \frac{\partial \mathbf{F}}{\partial R_2^k} = \begin{bmatrix} \frac{s_1 CR_1 (1+s_1 CR_1)}{[1+s_1 C(R_1+R_2^k)]^2} \\ : \\ \frac{s_m CR_1 (1+s_m CR_1)}{[1+s_m C(R_1+R_2^k)]^2} \end{bmatrix} \quad (6.17)$$

is a column vector, it has a rank of one (full rank) if nonzero frequencies are used. Also, from (6.15) it is quite obvious that

$$\mathbf{J}_c = \begin{bmatrix} \frac{\partial \mathbf{F}}{\partial R_1} & \frac{\partial \mathbf{F}}{\partial C} \end{bmatrix} \quad (6.18)$$

is a function of  $R_2$ , hence the second necessary condition is also satisfied. Bandler, Chen and Daijavad (1986b) have shown that by using two different frequencies a full rank Jacobian  $G$  can be obtained, hence  $R_1^0$ ,  $C^0$ ,  $R_2^0$  and  $R_2^1$  can be uniquely identified from the measurements on  $V_2$ . This simple example has demonstrated the potential of improving identifiability by using multiple circuits.

In practice, especially for a complicated circuit, we may not observe a clean-cut deficiency in identifiability. Still, difficulties may arise from an ill-conditioned Jacobian. In such cases we may be able to formulate a better conditioned problem by the multi-circuit approach.

#### 6.4.2 Multi-Circuit Modeling of a FET Device

The development of an equivalent circuit model for FET devices is an important part of the design process of monolithic microwave integrated circuits (MMIC). The small-signal equivalent circuit model shown in Fig. 6.4 is widely used in the literature

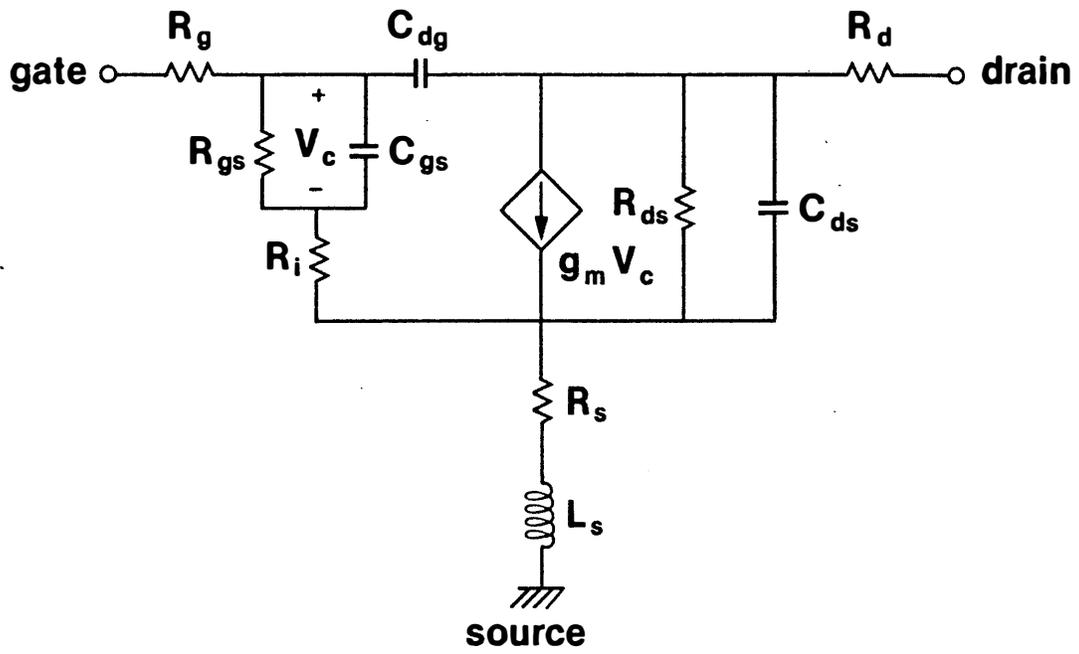


Fig. 6.4 Small-signal equivalent circuit model for a FET device.

(e.g., Curtice and Camisa 1984) and by commercial packages like TOUCHSTONE (1985) and SUPER-COMPACT (1986). Practical modeling of FET devices has been frequently troubled by nonunique solutions.

To demonstrate the multi-circuit modeling approach, we utilize three sets of actual measurements on scattering parameters of a FET device which were taken at 17 frequency points from 2GHz to 18GHz, 1GHz apart, under the following biasing conditions (Pucel 1986).

1.  $V_{ds} = 4V$ ,  $V_{gs} = 0.00V$ ,  $I_{ds} = 177mA$ .
2.  $V_{ds} = 4V$ ,  $V_{gs} = -1.74V$ ,  $I_{ds} = 92mA$ .
3.  $V_{ds} = 4V$ ,  $V_{gs} = -3.10V$ ,  $I_{ds} = 37mA$ .

Eleven model parameters, namely

$$\{R_g, R_d, L_s, \tau, R_{ds}, R_i, R_s, C_{gs}, C_{dg}, C_{ds}, g_m\},$$

are taken as variables. The first four parameters are considered to be bias insensitive and, therefore, treated as common variables.

Following formulas (6.3) and (6.4), we have

$$\mathbf{x} = \begin{bmatrix} \phi^0 \\ \phi_a^1 \\ \phi_a^2 \end{bmatrix}. \quad (6.19)$$

The vector of reference parameters actually has two parts as

$$\phi^0 = \begin{bmatrix} \phi_c \\ \phi_a^0 \end{bmatrix}, \quad (6.20)$$

where  $\phi_c$  consists of the common variables as

$$\phi_c = [R_g \ R_d \ L_s \ \tau]^T. \quad (6.21)$$

Also, for  $k = 0, 1, 2$ , we have

$$\phi_a^k = [R_{ds}^k \ R_i^k \ R_s^k \ C_{gs}^k \ C_{dg}^k \ C_{ds}^k \ g_m^k]^T. \quad (6.22)$$

The total number of variables is 25.

For each circuit and each frequency, by treating the real and imaginary parts of the four scattering parameters separately, we created 8 discrete error functions. Overall, a total of 408 error functions was considered.

The superlinearly convergent  $\ell_1$  algorithm described by Bandler, Kellermann and Madsen (1987) was utilized to solve the resulting optimization problem. The identified model parameters are given in Table 6.1. The match between the model responses and the measurements, at both the starting point and the solution, is shown in Figs. 6.5, 6.6 and 6.7. Here, the exact sensitivities of the scattering parameters as required by the optimization algorithm were calculated following the approach described in Chapter 3. In Chapter 7 we will solve the same problem without requiring exact sensitivities but, instead, utilizing approximate gradients.

#### 6.4.3 Model Verification for a Multi-Coupled Cavity Filter

The multi-circuit approach can also be applied to model verification. This is typically related to cases where the parasitic uncertainty  $\Delta F^L$  (Section 6.3.1) has put the validity of a model in doubt. Instead of explicitly defining common and independent variables, we use the formulation given by (6.7) and (6.8), where all the parameters are allowed to vary and the deviation of each parameter from its reference value forms a penalty term in the  $\ell_1$  objective function. By this formulation, our confidence in the model is strengthened if the result not only produces a reasonable

TABLE 6.1  
PARAMETER VALUES OF THE FET MODELS

Parameter	Starting Point	Solution		
		Circuit 1	Circuit 2	Circuit 3
$R_g$ (OH)	1.0	2.6025	2.6025	2.6025
$R_d$ (OH)	1.0	3.7630	3.7630	3.7630
$R_{ds}$ (OH)	143.0	199.1591	163.8249	163.1911
$R_i$ (OH)	1.0	0.0099	0.0999	0.3891
$R_s$ (OH)	1.0	1.0016	0.9220	0.6482
$L_s$ (nH)	0.02	0.0039	0.0039	0.0039
$C_{gs}$ (pF)	1.4	0.7181	0.4417	0.3454
$C_{dg}$ (pF)	0.07	0.0306	0.0475	0.0609
$C_{ds}$ (pF)	0.4	0.2228	0.2229	0.2151
$g_m$ (/OH)	0.09	0.0696	0.0521	0.0410
$\tau$ (ps)	7.0	3.9558	3.9558	3.9558

Biasing Conditions

Circuit 1:	$V_{ds}=4V$	$V_{gs}= 0.00V$	$I_{ds}=177mA$
Circuit 2:	$V_{ds}=4V$	$V_{gs}=-1.74V$	$I_{ds}= 92mA$
Circuit 3:	$V_{ds}=4V$	$V_{gs}=-3.10V$	$I_{ds}= 37mA$

The three circuits start with identical model parameter values.

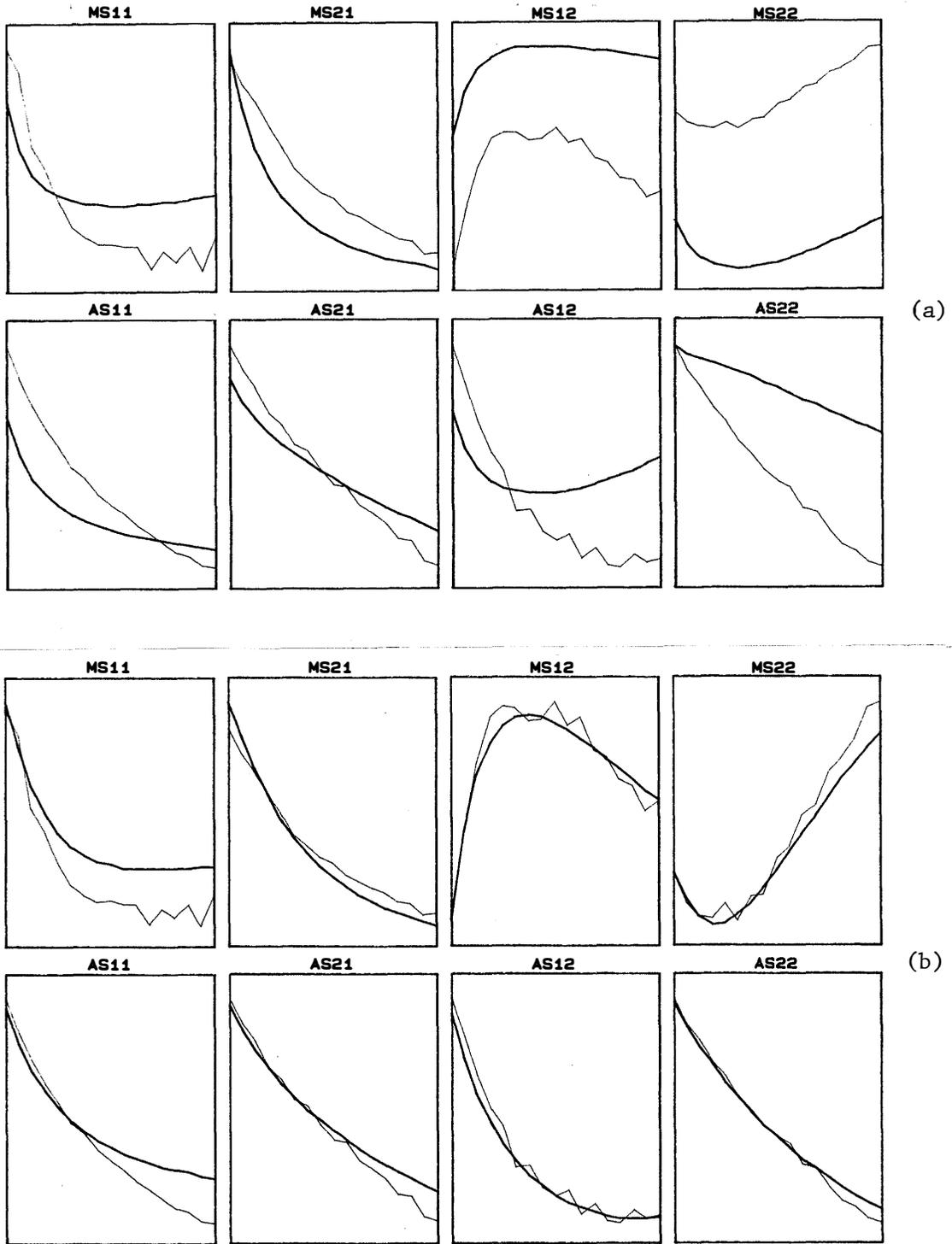


Fig. 6.5 The scattering parameter match between the FET model and the measurements at (a) the starting point and (b) the solution, with the biasing conditions being  $V_{ds} = 4V$ ,  $V_{gs} = 0V$  and  $I_{ds} = 177mA$ .

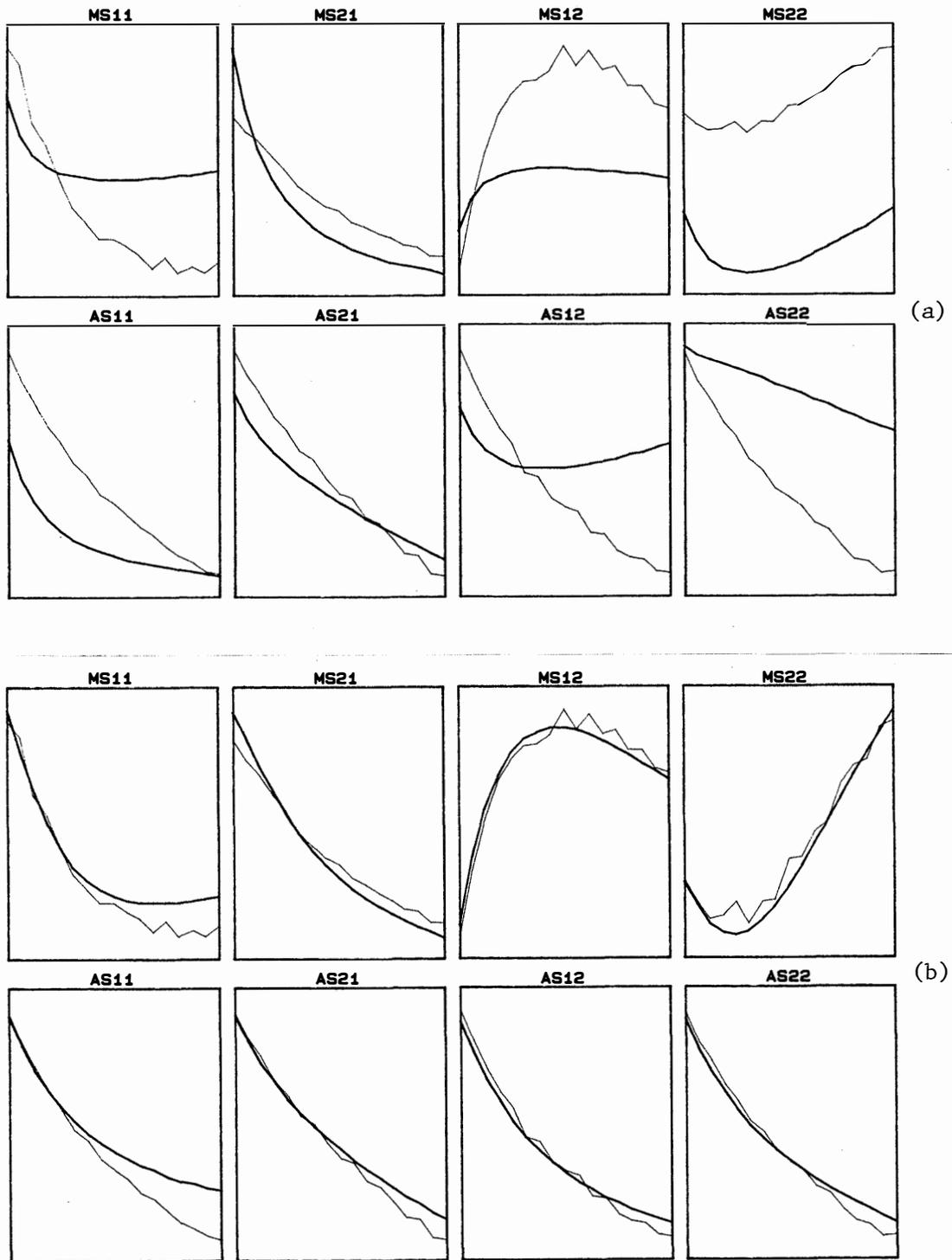


Fig. 6.6 The scattering parameter match between the FET model and the measurements at (a) the starting point and (b) the solution, with the biasing conditions being  $V_{ds} = 4V$ ,  $V_{gs} = -1.74V$  and  $I_{ds} = 92mA$ .

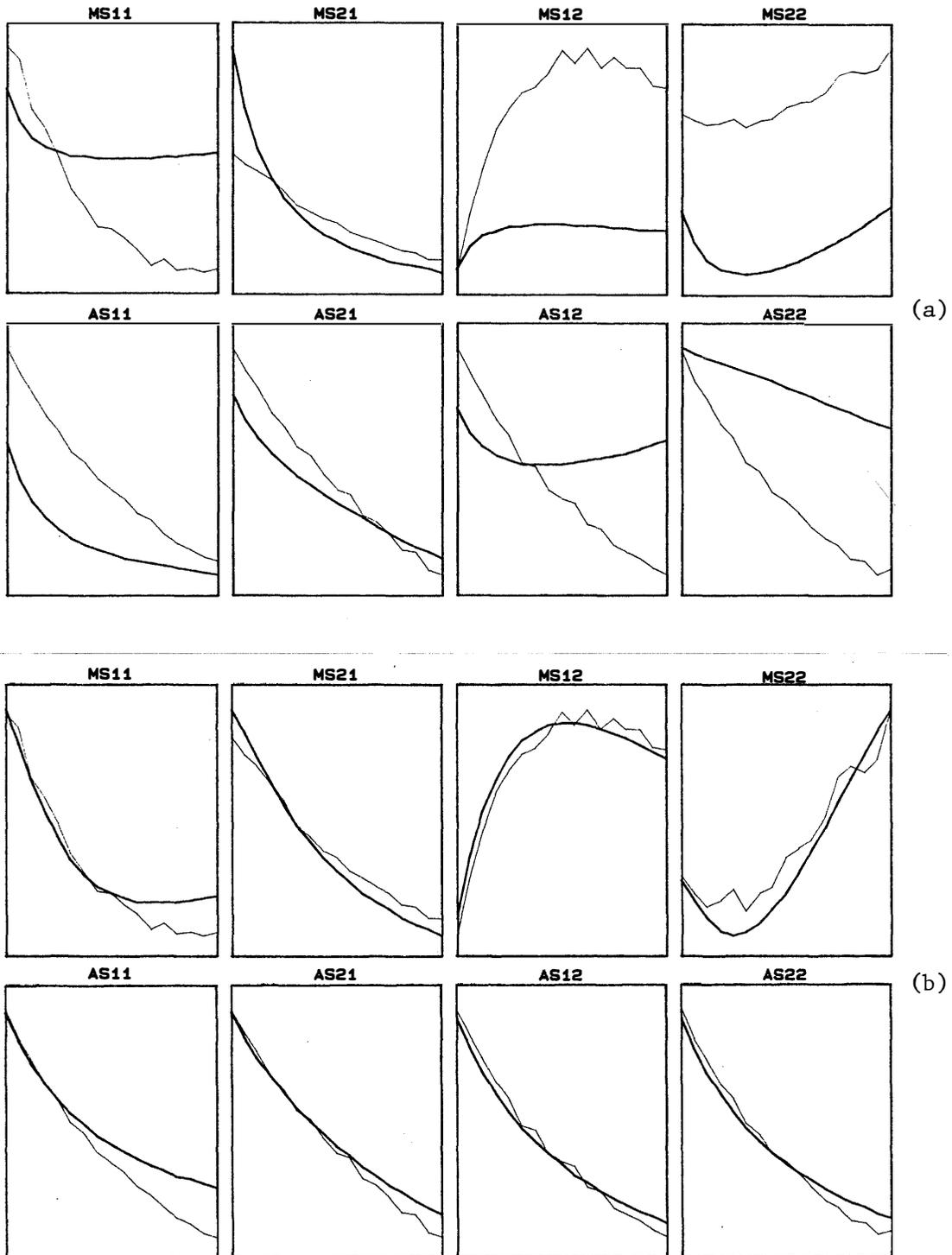


Fig. 6.7 The scattering parameter match between the FET model and the measurements at (a) the starting point and (b) the solution, with the biasing conditions being  $V_{ds} = 4V$ ,  $V_{gs} = -3.1V$  and  $I_{ds} = 37mA$ .

match between the measured and calculated responses, but also demonstrates consistency with respect to the physical adjustments. Otherwise we should probably reject the current model and consider a more adequate one.

As a practical example, consider an 8th-order multi-coupled cavity filter centered at 11902.5MHz with a 60MHz bandwidth. The general structure and equivalent circuit model of these filters have been described in Chapter 4. The return loss and insertion loss measurements of an optimally tuned filter and the same filter after an adjustment on the iris which dominantly controls coupling  $M_{23}$  were provided by ComDev Ltd. (1985). Based on the physical structure of the filter, screw couplings  $M_{12}$ ,  $M_{34}$ ,  $M_{56}$  and  $M_{78}$  and iris couplings  $M_{23}$ ,  $M_{45}$ ,  $M_{67}$  and  $M_{58}$ , as well as all cavity resonant frequencies and input/output transformer ratios are considered as possible nonzero parameters to be identified.

In the first attempt, the stray coupling  $M_{36}$  (a parasitic element) was ignored. The parameters of the equivalent circuit model, which did not include  $M_{36}$ , were identified from the measurements and summarized in Table 6.2. An examination of the results shows no apparent trend for the change in parameters, i.e., it would have been impossible to guess the source of perturbation (an adjustment on the iris controlling  $M_{23}$ ) from these results. This kind of inconsistency would not have been discovered if only a single circuit had been considered.

Consequently, the inconsistent model was rejected. A new model which included the stray coupling  $M_{36}$  was introduced and we

TABLE 6.2  
IDENTIFIED PARAMETER VALUES FOR THE 8TH-ORDER FILTER

Coupling	$M_{36}$ ignored		$M_{36}$ present	
	Original	Perturbed	Original	Perturbed
$M_{11}$	-0.0306	-0.1122	-0.0260	-0.0529
$M_{22}$	0.0026	-0.0243	0.0354	0.6503*
$M_{33}$	-0.0176	-0.0339	-0.0674	-0.6113*
$M_{44}$	-0.0105	-0.0579	-0.0078	-0.0151
$M_{55}$	-0.0273	-0.0009	-0.0214	0.0506
$M_{66}$	-0.0256	0.0457	-0.0179	-0.0027
$M_{77}$	-0.0502	0.0679	-0.0424	-0.0278
$M_{88}$	-0.0423	0.0594	-0.0426	-0.0272
$M_{12}$	0.7789	0.7462	0.3879	0.2876*
$M_{23}$	0.8061	0.8376	0.9990	0.8160*
$M_{34}$	0.4460	0.4205	0.0270	-0.1250*
$M_{45}$	0.5335	0.5343	0.4791	0.5105
$M_{56}$	0.5131	0.5373	0.5006	0.5026
$M_{67}$	0.7260	0.7469	0.6495	0.6451
$M_{78}$	0.8330	0.8476	0.8447	0.8463
$M_{14}$	0.3470	-0.3582	-0.7648	-0.7959
$M_{58}$	-0.1995	-0.1892	-0.1000	-0.0953
$M_{36}$	-	-	0.1314	0.1459

Input and output couplings:  $n_1^2 = n_2^2 = 1.067$

\* Significant change in parameter value.

processed exactly the same measurements as before. The filter parameters identified in this case are also contained in Table 6.2. A comparison of the original and perturbed filter parameters shows significant changes in  $M_{12}$ ,  $M_{23}$ ,  $M_{34}$ ,  $M_{22}$  and  $M_{33}$  (all related to cavities 2 and 3), which is absolutely consistent with the actual adjustment. By inspecting the change in model parameters, it is possible to deduce which physical parameter has been adjusted. The match between the model responses and the measurements is shown in Figs. 6.8 and 6.9.

It is worth mentioning that the calculated responses of the inconsistent model, which ignored  $M_{36}$ , match the measured responses almost as well as the correct model. This justifies the essence of the multi-circuit approach, which attempts to identify the most consistent set of parameters among many that produce a reasonable and similar match between measured and calculated responses.

#### 6.4.4 Modeling of the Relationship Between Physical and Equivalent Circuit Parameters

Another important application of multi-circuit modeling is to create analytical formulas which relate the model  $\phi$  to the physical parameters  $\phi^M$ . Such formulas will become extremely useful in guiding an actual production alignment or postproduction tuning procedure.

A sequence of adjustments on  $\phi^M$  can be systematically made and multiple sets of measurements can be taken. By single-circuit optimization, these measurements would be processed separately to

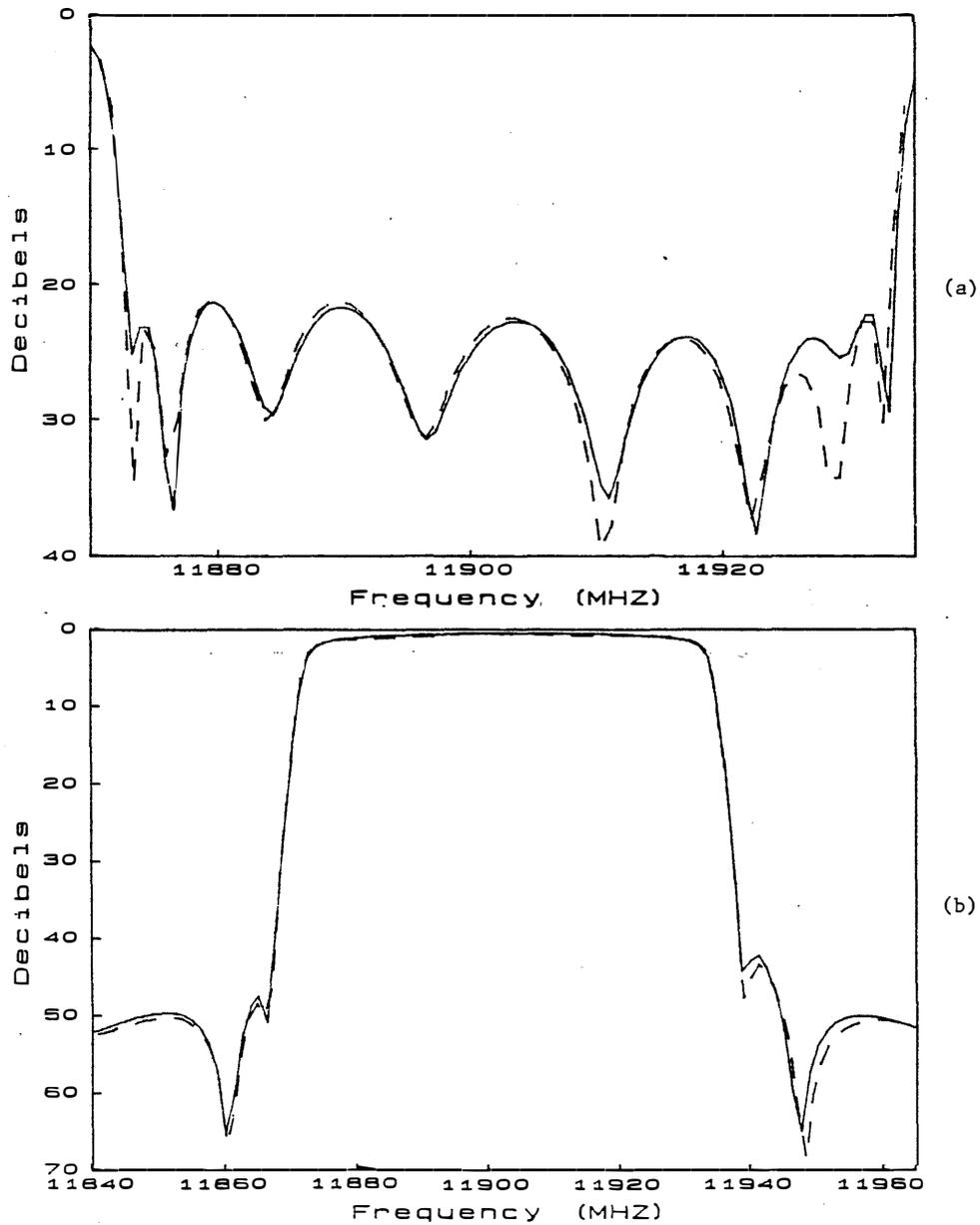


Fig. 6.8 (a) Input return loss and (b) insertion loss of the 8th-order filter before adjusting the iris. The solid line represents the model response and the dashed line shows the measured response.

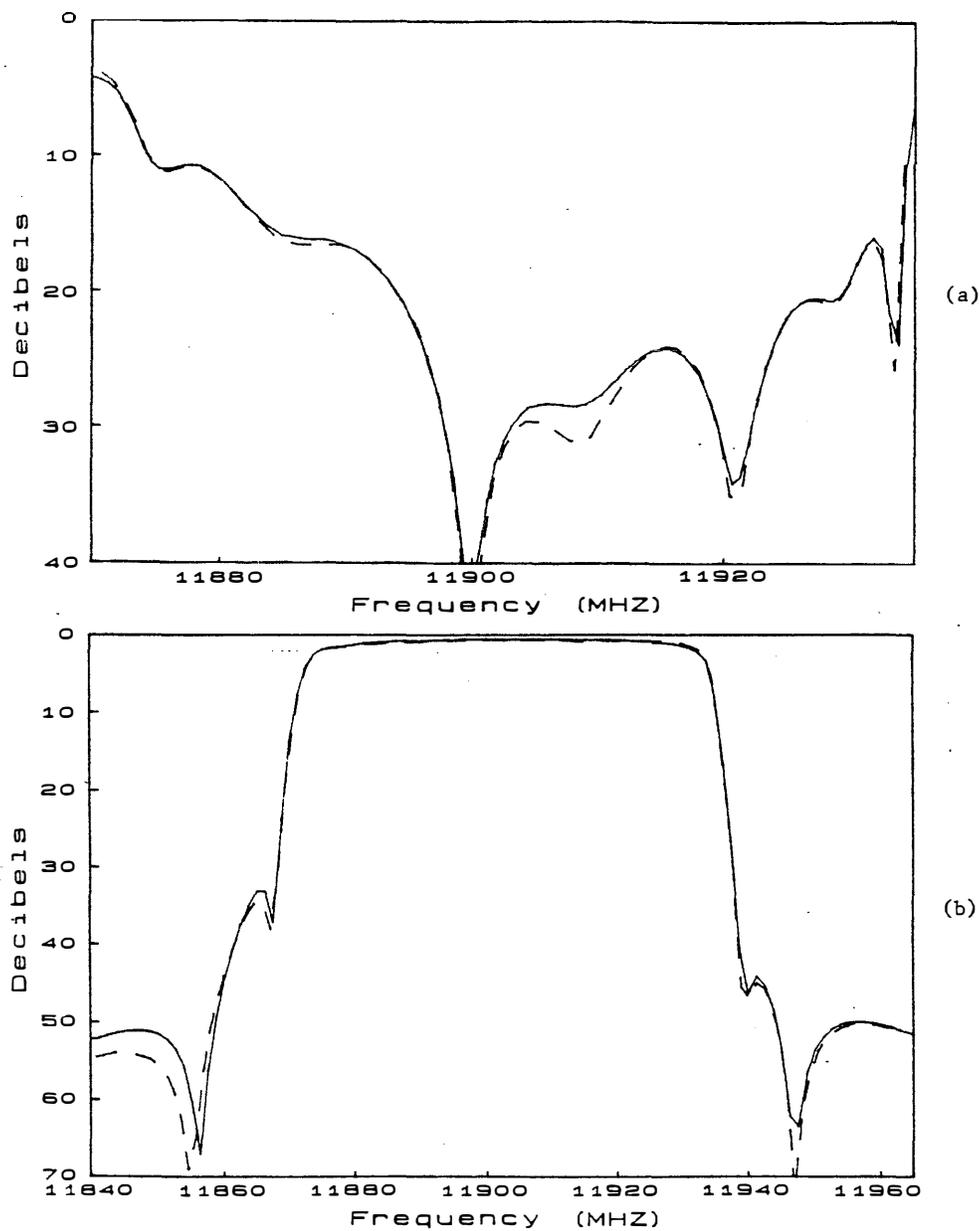


Fig. 6.9 (a) Input return loss and (b) insertion loss of the 8th-order filter after adjusting the iris. The solid line represents the model response and the dashed line shows the measured response.

obtain a set of static models. In the presence of uncertainties, a single change in  $\phi^M$  may seem to cause fluctuations in all the model parameters. Obviously, such results are of very little use. By including penalty terms in the objective function (6.8) and using the  $\ell_1$  optimization, we may be able to suppress small fluctuations in the model parameter values and emphasize the dominant relationships. In other words, the multi-circuit approach is more likely to produce models that are meaningful and useful in practice. Simply put, it certainly makes sense to process simultaneously the measurements that are made systematically.

Actually, the variables to be optimized need not always be the model parameters. They can as well include the coefficients of a proposed formula.

As an example, consider a 6th-order multi-coupled cavity filter centered at 11783MHz with a 56MHz bandwidth. Three coupling screws, whose positions are represented as elements of  $\phi^M$ , were adjusted. These screws were assumed to control the couplings  $M_{12}$ ,  $M_{34}$  and  $M_{56}$ , which are model parameters in  $\phi$ . Starting from a reference position, each screw was adjusted four times, twice in the clockwise direction (screw increasingly penetrates the cavity) for 90 and 180 degrees, as well as twice counterclockwise. After each adjustment, filter responses (input and output return loss, insertion loss and group delay) were measured (the measurement data was furnished by ComDev Ltd. (1986)). Using the techniques we have described, the model parameters were identified.

The variation of the identified coupling values versus the

relative position of the screws is shown in Figs. 6.10, 6.11 and 6.12 (these figures were originally produced by Daijavad (1986)). The correlation between each screw and the coupling it dominantly controls is clearly pronounced while fluctuations in other parameters are kept minimal.

#### 6.5 CONCLUDING REMARKS

In this chapter, we have discussed under several categories uncertainties which tend to deteriorate the results of modeling. We have introduced a multi-circuit approach which, compared with the more traditional concept of nominal circuit formulation, is more likely to produce consistent and reliable results in the presence of various uncertainties.

The mathematical formulations of the new approach as well as some relevant computational considerations have been presented. The unique property of the  $\ell_1$  optimization in relation to modeling has been exposed. A suitable measure of identifiability has been defined and two necessary conditions for improving a rank-deficient system by using multiple circuits have been developed. Practical applications to the unique identification of model parameters, model verification and tuning-related modeling have been described in detail. These applications have been illustrated and justified through significant examples including the modeling of a FET device and multi-coupled cavity filters.

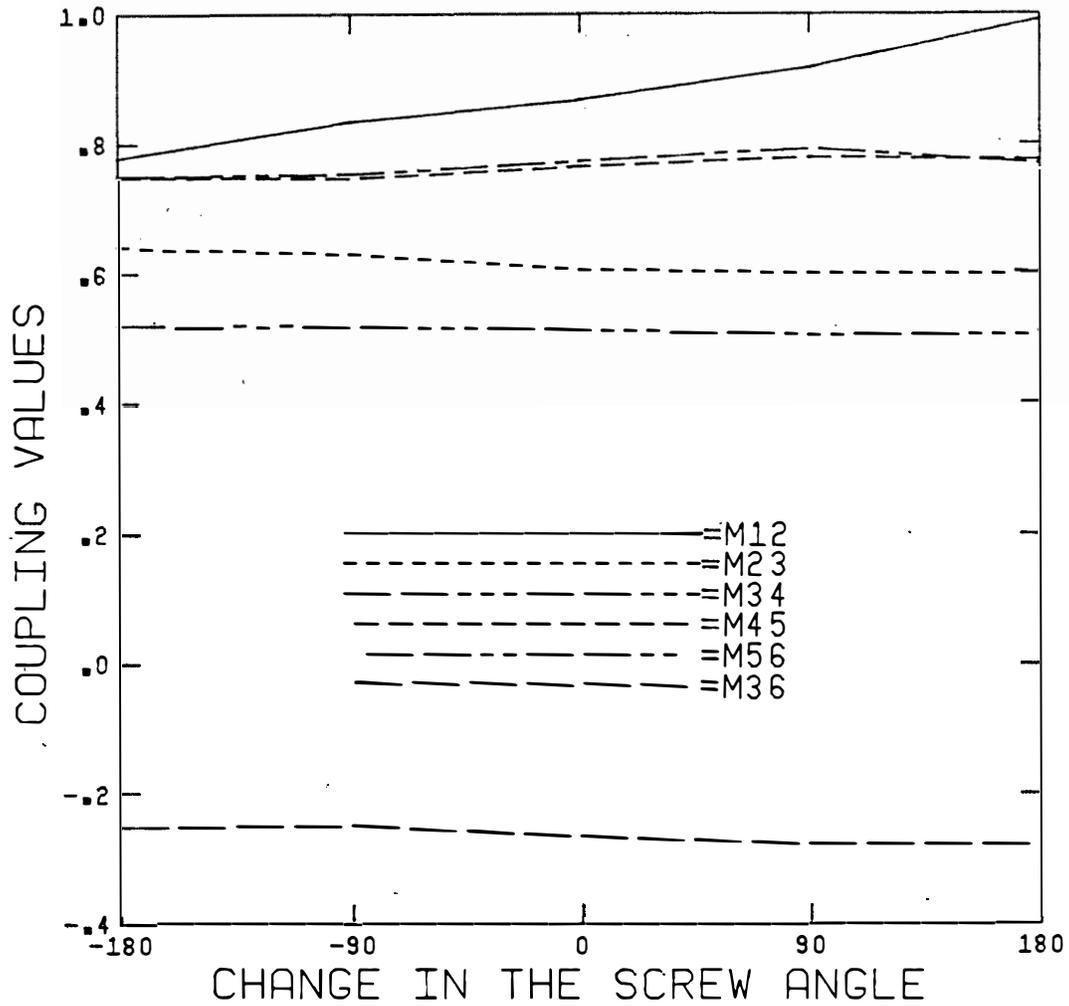


Fig. 6.10 The identified coupling values versus the relative position in degrees of the screw which is assumed to control  $M_{12}$  dominantly.

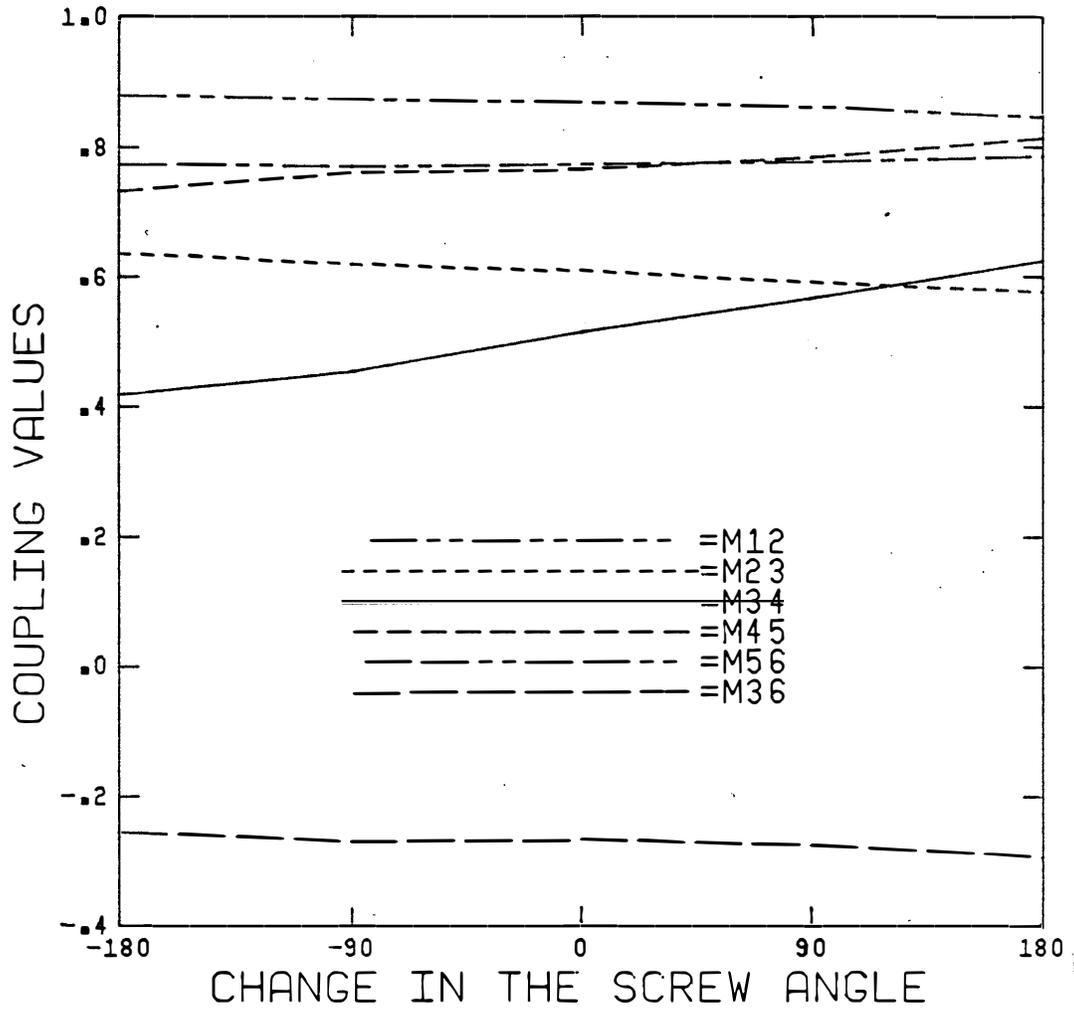


Fig. 6.11 The identified coupling values versus the relative position in degrees of the screw which is assumed to control  $M_{34}$  dominantly.

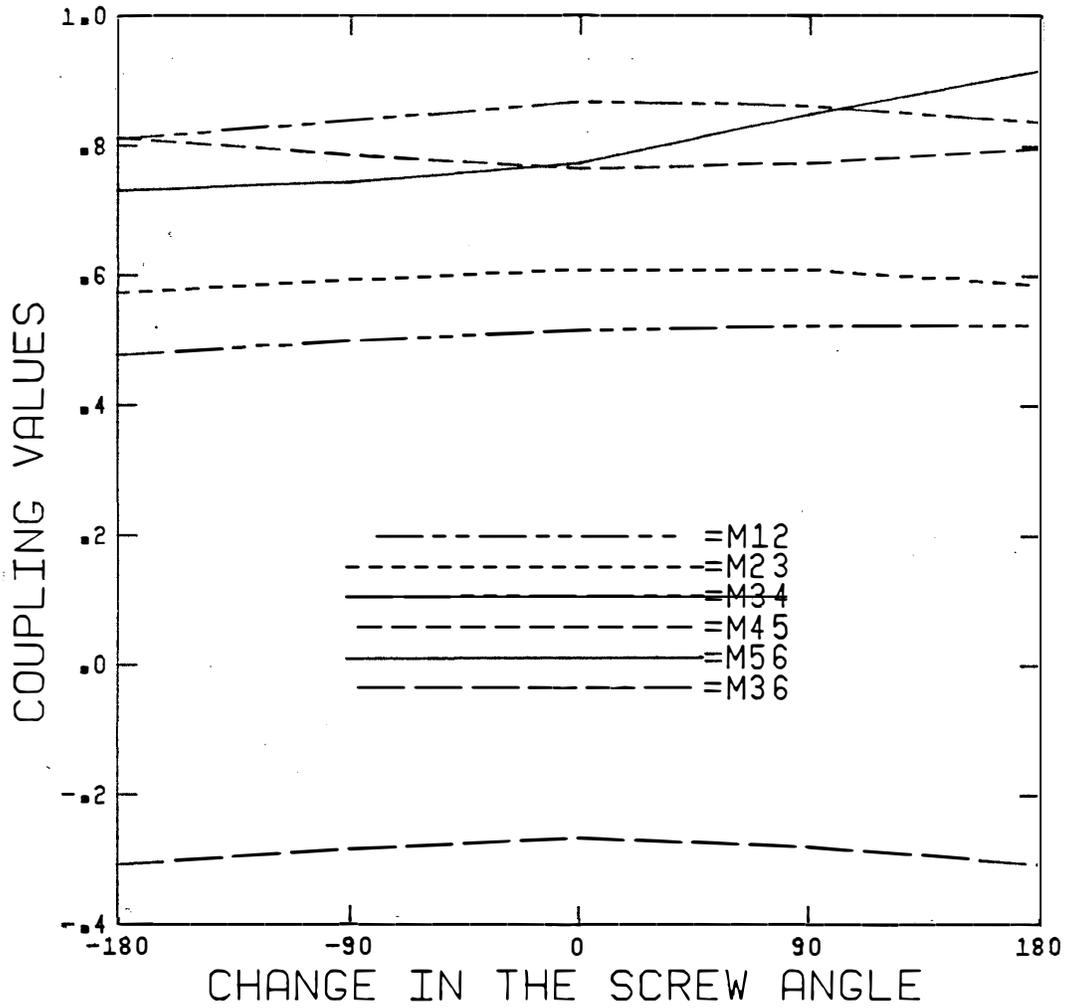


Fig. 6.12 The identified coupling values versus the relative position in degrees of the screw which is assumed to control  $M_{56}$  dominantly.

## CHAPTER 7

### OPTIMIZATION WITH INTEGRATED GRADIENT APPROXIMATIONS

#### 7.1 INTRODUCTION

Many powerful gradient-based algorithms have been developed in recent years for nonlinear optimization. However, the effort to extend their application to a wide range of practical problems has been frustrated by the requirement of exact gradients of all functions with respect to all variables. For some applications, either an explicit sensitivity expression is not available, e.g., when time-domain analysis and nonlinear circuits are involved, or the actual evaluation of such an expression is very tedious and time-consuming, e.g., for large-scale networks. Partly due to these difficulties, exact sensitivity calculations have yet to be implemented in many general-purpose CAD software packages, although the concept of the adjoint network has been in existence for nearly two decades and has had success in many specialized applications. The inability or inconvenience in calculating the exact derivatives has created a gap between the theoretical advances in gradient-based nonlinear optimization techniques and their actual implementation.

With only the function values available, as is the case for many CAD packages on the market, one usually resorts to the method of perturbations (finite differences) for gradients. However, except for rather simple problems, the computational labor for estimating gradients entirely by perturbations is very expensive.

This chapter addresses itself to a flexible and effective approach to optimization with integrated gradient approximations (Bandler, Chen, Daijavad and Madsen, 1986, 1987). It is a hybrid approach which incorporates the use of perturbations, the Broyden update (Broyden 1965) and the special iterations of Powell (1970a).

Approximations to second-order derivatives have been extensively studied in the context of quasi-Newton methods. Those results are not directly applicable to gradient approximations, because certain important properties of a Hessian such as symmetry and positive definiteness are not generally relevant to a Jacobian.

Previous work on gradient approximations has been reported by Madsen (1975) and Zuberek (1984). They have used the Broyden rank-one formula in conjunction with the special iterations of Powell in their work. Such an implementation may not be able to provide sufficiently accurate results for highly nonlinear problems or for certain optimization techniques. In the approach described in this chapter, perturbations are integrated in a flexible manner to allow regular corrections to the approximate gradients. Therefore, a suitable compromise between accuracy and computational labor may be achieved for various applications. We also propose a modification of the Broyden update which incorporates a knowledge, if available, of the structure of the problem (e.g., one that has a sparse Jacobian).

Suitable methods of integrating gradient approximations with an optimization algorithm are also discussed in this chapter. A general-purpose transparent interface is described. Specific

examples of integrated optimization are provided through a minimax and an  $\ell_1$  implementation. The performance of these algorithms is tested using several standard problems. The practical usefulness of the new method is demonstrated through significant and diversified circuit applications, including examples of worst-case design, multiplexer optimization, fault location and FET modeling.

## 7.2 GRADIENT APPROXIMATIONS

### 7.2.1 Estimating Gradient by Perturbations

The first-order derivative of  $f_j(\mathbf{x})$  with respect to  $x_i$  can be estimated by

$$\frac{\partial f_j(\mathbf{x})}{\partial x_i} \approx \frac{f_j(\mathbf{x} + h\mathbf{u}_i) - f_j(\mathbf{x})}{h}, \quad (7.1)$$

where  $\mathbf{u}_i$  is a column vector which has 1 in the  $i$ th position and zeros elsewhere, as has been consistently used throughout this thesis. The accuracy of such an estimate may be improved by using a smaller  $h$  as well as by averaging the results of a two-sided approximation (using both positive and negative perturbations). This method has been widely used by commercial packages such as TOUCHSTONE (1985) and SUPER-COMPACT (1986), since it is straight forward and quite reliable. However, the computational labor involved grows in proportion to the dimension of the problem.

In the new algorithm described in this chapter, perturbations are used to obtain an initial approximation to the gradient at the starting point of an optimization process, unless such an initial approximation is already available (e.g., it may have been

stored from a previous optimization) and can be retrieved. During the optimization, we may also incorporate a regular use of perturbations to maintain the accuracy of gradient approximations at a desirable level.

### 7.2.2 The Broyden Update

The Broyden update refers to a rank-one formula proposed by Broyden (1965) as

$$\mathbf{G}_{k+1} = \mathbf{G}_k + \frac{f(\mathbf{x}_k + \mathbf{h}_k) - f(\mathbf{x}_k) - \mathbf{G}_k \mathbf{h}_k}{\mathbf{h}_k^T \mathbf{h}_k} \mathbf{h}_k^T, \quad (7.2)$$

where  $\mathbf{G}_k$  is an approximation of the Jacobian  $[\partial f^T / \partial \mathbf{x}]^T$  at  $\mathbf{x}_k$ ,  $\mathbf{h}_k$  is an increment vector and  $\mathbf{G}_{k+1}$  provides an updated Jacobian. The values of the function  $f$  at  $\mathbf{x}_k$  and  $(\mathbf{x}_k + \mathbf{h}_k)$  are assumed available. If these two points ( $\mathbf{x}_k$  and  $(\mathbf{x}_k + \mathbf{h}_k)$ ) are iterates of the optimization algorithm, then the Broyden update requires no additional function evaluations, regardless of the dimension of the problem.

Apparently, the approximate Jacobians generated by the Broyden update are in general less accurate as compared with those obtained from perturbations. Hence, the optimization may require more steps to reach the solution or may not reach the correct solution at all. Broyden (1965) has shown that for quadratic functions the Broyden update will converge and will reduce the overall computational effort. Although such properties can not be proved for a general nonlinear problem, the Broyden update still provides an efficient alternative for approximating derivatives.

From (7.2), we can verify that the updated approximation

$G_{k+1}$  satisfies the following equation

$$f(\mathbf{x}_k + \mathbf{h}_k) - f(\mathbf{x}_k) = G_{k+1} \mathbf{h}_k . \quad (7.3)$$

In other words,  $G_{k+1}$  provides a perfect linear interpolation between the two points  $\mathbf{x}_k$  and  $(\mathbf{x}_k + \mathbf{h}_k)$ .

Some difficulties in the application of the Broyden update have been observed by many researchers (Powell 1970a, Madsen 1975 and Zuberek 1984).

(1) If some functions are linear in some variables and if the corresponding components of  $\mathbf{h}_k$  are nonzero, then the approximation of constant derivatives are updated by nonzero values. We illustrate this difficulty by a simple example. Let  $f_j = x_1^2 + 2x_3$  be a function in  $f$ . Denote the variables by  $\mathbf{x} = [x_1 \ x_2 \ x_3]^T$  and the gradient by  $\mathbf{f}'_j(\mathbf{x}) = [g_1 \ g_2 \ g_3]^T$ . Two components of the gradient, namely  $g_2 = 0$  and  $g_3 = 2$ , are constants and can be found accurately by perturbations.  $g_1$  is the only component that needs to be updated. Suppose that  $\mathbf{x}_k = [1 \ 1 \ 1]^T$ ,  $\mathbf{h}_k = [0.5 \ 0.5 \ 0.5]^T$  and a perfect estimation of  $\mathbf{f}'_j(\mathbf{x}_k)$  is available as  $[2 \ 0 \ 2]^T$ . The approximation to  $\mathbf{f}'_j(\mathbf{x}_k + \mathbf{h}_k)$ , as given by the Broyden update, would be  $[2.167 \ 0.167 \ 2.167]^T$  (the true value is  $[3 \ 0 \ 2]^T$ ).

(2) Along directions orthogonal to  $\mathbf{h}_k$  the Jacobian is not updated. Mathematically it can be verified from (7.2) that

$$G_{k+1} \mathbf{p} = G_k \mathbf{p}, \quad \text{for } \mathbf{p}^T \mathbf{h}_k = 0. \quad (7.4)$$

To overcome these difficulties, we implement a weighted updating formula and the special iterations of Powell (1970a).

### 7.2.3 A Weighted Broyden Update

The weighted update is to be applied to the Jacobian matrix on a row-by-row basis, i.e., we update the gradient vectors of individual functions. The  $j$ th row vector of the approximate Jacobian, denoted by  $(\mathbf{g}_j)_k$ , is an approximation to  $\mathbf{f}'_j(\mathbf{x}_k)$ , the gradient of  $f_j$ .

Suppose that the Hessian of  $f_j$  is available to us and denoted by  $H_j$ , then

$$\mathbf{f}'_j(\mathbf{x}_k + \mathbf{h}_k) \approx \mathbf{f}'_j(\mathbf{x}_k) + H_j(\mathbf{x}_k) \mathbf{h}_k. \quad (7.5)$$

Analogously to (7.5), we devise an updating formula to obtain an approximation to  $\mathbf{f}'_j(\mathbf{x}_k + \mathbf{h}_k)$  as

$$(\mathbf{g}_j)_{k+1} = (\mathbf{g}_j)_k + \alpha H_j(\mathbf{x}_k) \mathbf{h}_k. \quad (7.6)$$

If we choose the coefficient  $\alpha$  as

$$\alpha = \frac{f_j(\mathbf{x}_k + \mathbf{h}_k) - f_j(\mathbf{x}_k) - (\mathbf{g}_j)_k^T \mathbf{h}_k}{\mathbf{h}_k^T H_j(\mathbf{x}_k) \mathbf{h}_k}, \quad (7.7)$$

then the linear model as given by (7.3) will be preserved, namely

$$f_j(\mathbf{x}_k + \mathbf{h}_k) - f_j(\mathbf{x}_k) = (\mathbf{g}_j)_{k+1}^T \mathbf{h}_k. \quad (7.8)$$

In practice we are very unlikely to have access to the Hessian of any  $f_j$ . Even so, two basic facts are obvious: the Hessian of a quadratic function is constant, and if  $f_j$  is linear in  $x_i$  then the  $i$ th row and the  $i$ th column of the Hessian contain only zeros. Hence, we propose the use of a constant diagonal matrix

$$\mathbf{W}_j = \text{diag}[w_{j1} \dots w_{jn}], \quad w_{ji} \geq 0, \quad i = 1, \dots, n, \quad (7.9)$$

where  $n$  is the dimension of  $\mathbf{x}$ . This leads to a weighted Broyden update as follows.

$$(\mathbf{g}_j)_{k+1} = (\mathbf{g}_j)_k + \frac{f_j(\mathbf{x}_k + \mathbf{h}_k) - f_j(\mathbf{x}_k) - (\mathbf{g}_j)_k^T \mathbf{h}_k}{\mathbf{q}_{jk}^T \mathbf{h}_k} \mathbf{q}_{jk} , \quad (7.10)$$

$$\mathbf{q}_{jk} = \mathbf{W}_j \mathbf{h}_k = [w_{j1}h_{k1} \ \dots \ w_{jn}h_{kn}]^T .$$

The weights  $w_{ji}$  provide a measure of the linearity of  $f_j$ . If  $f_j$  is linear in  $x_i$ , we set  $w_{ji}=0$ , and if  $f_j$  is nearly linear in  $x_i$ , we assign a small value to  $w_{ji}$ . It should be clear from (7.10) that only the relative magnitude of the weights is important, not their absolute values.

Consider the simple example we have used in Section 7.2.2, namely  $f_j = x_1^2 + 2x_3$ . Since  $f_j$  is independent of  $x_2$  and linear in  $x_3$ , we set  $w_{j2} = w_{j3} = 0$  and  $w_{j1} = 1$ . Under the same conditions as in Section 7.2.2, we obtain an approximate gradient using (7.10) as  $[2.5 \ 0 \ 2]^T$ , compared to the result given by the Broyden update as  $[2.167 \ 0.167 \ 2.167]^T$ , and the true gradient  $[3 \ 0 \ 2]^T$ .

The assignment of weights requires some knowledge of the functional relationship of  $f_j(\mathbf{x})$ . Such a knowledge may come from experience or may be gained from sensitivity analyses by performing a few perturbations. For instance, for a particular circuit, it may be known that some designable parameters have little influence on the performance function over some frequency or time intervals. Using an adaptive method to find  $\mathbf{W}_j$  might be of some theoretical interest. But it was felt to be unnecessary and too complicated to be practical at the present time.

The applications of the weighted update to practical circuit optimization are demonstrated later in this chapter.

#### 7.2.4 The Special Iterations of Powell

As has been shown in (7.4), along directions orthogonal to  $\mathbf{h}_k$  the approximate Jacobian is not updated by the Broyden formula. If some consecutive steps of optimization happen to be collinear, the updating procedure may not converge. Powell (1970a) suggested a method which produces strictly linearly independent directions. For this purpose, special iterations are introduced which intervene between the ordinary iterations of optimization. The increment vector of such a special iteration is not calculated to minimize the error functions, instead it serves the purpose of improving the accuracy of gradient approximations.

The algorithm for computing the increment vector for a special iteration, as derived by Powell (1970a), is given in the Appendix. Powell has shown that the application of the Broyden update on these specially generated directions is likely to improve the accuracy of derivative approximations.

#### 7.2.5 A Hybrid Approximation Algorithm

Our hybrid algorithm for gradient approximations consists of an initial approximation, the Broyden update, Powell's special iterations and regular corrections provided by perturbations.

At the starting point of optimization, the initial approximate Jacobian  $G_0$  is usually computed by perturbations. However,  $G_0$  may be already available, for example, it may have been stored from a previous optimization, and can be utilized to avoid unnecessary computations. This option would be useful if similar problems are

being solved repetitively (e.g., the same circuit is optimized with respect to different specifications). The accuracy of  $G_0$  is not very critical to the overall approximation. We have observed for some examples that convergence was achieved despite the erroneous estimates of  $G_0$ .

There is little hard evidence as to how frequently the special iterations should be used. Numerical experience, ours as well as other authors', has suggested the use of a special iteration between every two ordinary ones (i.e., every third iteration is a special iteration). Also, in our implementation, a special iteration is skipped provided that the changes in the functions agree fairly well with the linear prediction by the approximate gradient. This is considered to be true if

$$\|f_j(\mathbf{x}_k + \mathbf{h}_k) - f_j(\mathbf{x}_k) - G_k \mathbf{h}_k\| < 0.1 \|f_j(\mathbf{x}_k + \mathbf{h}_k) - f_j(\mathbf{x}_k)\|. \quad (7.11)$$

The purpose of this provision is to avoid unnecessary computations.

Whether perturbations should be used during optimization depends on the application. For small or mildly nonlinear problems, the Broyden update may suffice. For large-scale problems, especially in circuit applications where highly nonlinear functions are involved, the correction provided by perturbations is likely to be necessary. We have incorporated in our algorithm the use of perturbations with prescribed regularity, say, at every  $k$ th optimization iteration.

The Broyden update with or without weights, depending on whether the necessary knowledge of  $f(\mathbf{x})$  is available, is employed between perturbations.

This hybrid approximation method has proved to be flexible, effective and efficient for a large variety of applications.

#### 7.2.6 Integration with Optimization Methods

Software for gradient-based optimization typically requires a user-defined routine which accepts a set of values for  $\mathbf{x}$  as input and returns the values of  $f(\mathbf{x})$  as well as the first-order derivatives.

We have implemented an interface which integrates gradient approximations with optimization. Taking a set of values for  $\mathbf{x}$  from an optimizer, it calls a user-defined routine for the function values, carries out necessary operations for gradient approximations, and then returns to the optimizer the values of  $f(\mathbf{x})$  as well as the approximate Jacobian. The interface is transparent to both the optimizer and the user-defined simulation routine. The optimizer is provided with the required gradients, and the user-defined routine (typically a circuit simulation module) works as if the optimizer did not require gradients.

Some sophisticated optimization algorithms have distinct stages of operations. Typically, one of the stages is to be employed near the solution to accelerate the rate of convergence, for which the accuracy of the approximate gradient may become critical. Hence, it is desirable to allow the use of different approximation schemes at different phases of optimization. This and the other options related to gradient approximations (e.g., the use of perturbations) should be selected by the user.

We illustrate the actual integration of our algorithm through two specific implementations for the minimax and the  $\ell_1$  problems, respectively.

### 7.3 A MINIMAX ALGORITHM USING APPROXIMATE GRADIENTS

#### 7.3.1 Implementation of a Fortran Subroutine

The gradient approximations described in this chapter have been integrated with the Hald and Madsen (1981) algorithm for the minimax problem as

$$\text{minimize}_{\mathbf{x}} \max_j \{f_j(\mathbf{x})\} \quad (7.12)$$

subject to possible linear constraints. As has been described in Chapter 2, the Hald and Madsen algorithm consists of two stages, Stage 1 being a trust region Gauss-Newton method which provides global convergence and Stage 2 a quasi-Newton method which is intended to achieve a fast rate of convergence near a solution.

The original algorithm requires a user-supplied subroutine which calculates the values of the error functions as well as their first-order derivatives. This is replaced in the new algorithm by an interface which organizes gradient approximations as outlined in Section 7.2.6. The user's subroutine is required to supply only the function values. Also, the user is allowed to specify, via common block parameters, whether an initial approximate Jacobian is already available and how frequently perturbations should be used during optimization. When Stage 2 of the algorithm is activated near the solution, it is desirable to acquire and maintain a higher accuracy of the approximate gradients. To this end, the user may

request intensified corrections be provided by perturbations in Stage 2 (at a different rate than Stage 1).

The complete Fortran implementation is included in the KMOS software library which has been described in Chapter 2.

### 7.3.2 Performance on Some Test Problems

A large variety of problems have been solved using the new algorithm. In this section we present the results for some test problems. Convergence was achieved for all the test problems. More precisely, the solutions obtained using approximate gradients agree with those using exact gradients to five significant figures. A comparison of computational effort between the new algorithm and a method that uses perturbations only is given in Table 7.1.

#### Problem M1

Consider the classical two-section 10:1 transmission-line transformer shown in Fig. 7.1. Originally proposed by Bandler and Macdonald (1969), this problem has been widely used to test minimax algorithms. The error functions ( $f_j$ ) are given by the reflection coefficient sampled at 11 frequencies normalized with respect to 1GHz: {0.5, 0.6, ..., 1.5}. Madsen and Schjaer-Jacobsen (1976) have shown that when we take the characteristic impedances  $Z_1$  and  $Z_2$  as variables and keep the lengths  $l_1$  and  $l_2$  constant at their optimal values (the quarter wavelength at the center frequency), the minimax problem is singular. To solve it effectively, the quasi-Newton iteration (Stage 2) of the algorithm is necessary.

Fig. 7.2 illustrates, on a minimax contour diagram, the

TABLE 7.1

COMPARISON OF COMPUTATIONAL EFFORT FOR THE MINIMAX EXAMPLES

Problem	Number of Function Evaluations	
	Entirely by Perturbations	By the New Algorithm
M1	24 (8)	18 (10)
M2	24 (8)	18 (12)
M3	59 (11)	30 (18)
M4	84 (11)	66 (41)
M5	9 (3)	5 (3)
M6	32 (10)	19 (14)
M7	29 (9)	14 (11)

The entries in parentheses are numbers of optimization iterations.

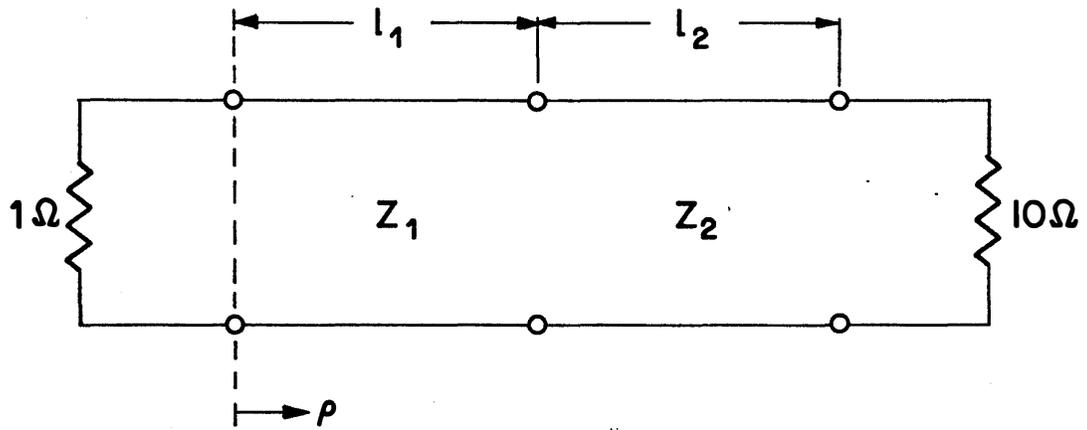


Fig. 7.1 Two-section, 10:1 transmission-line transformer.

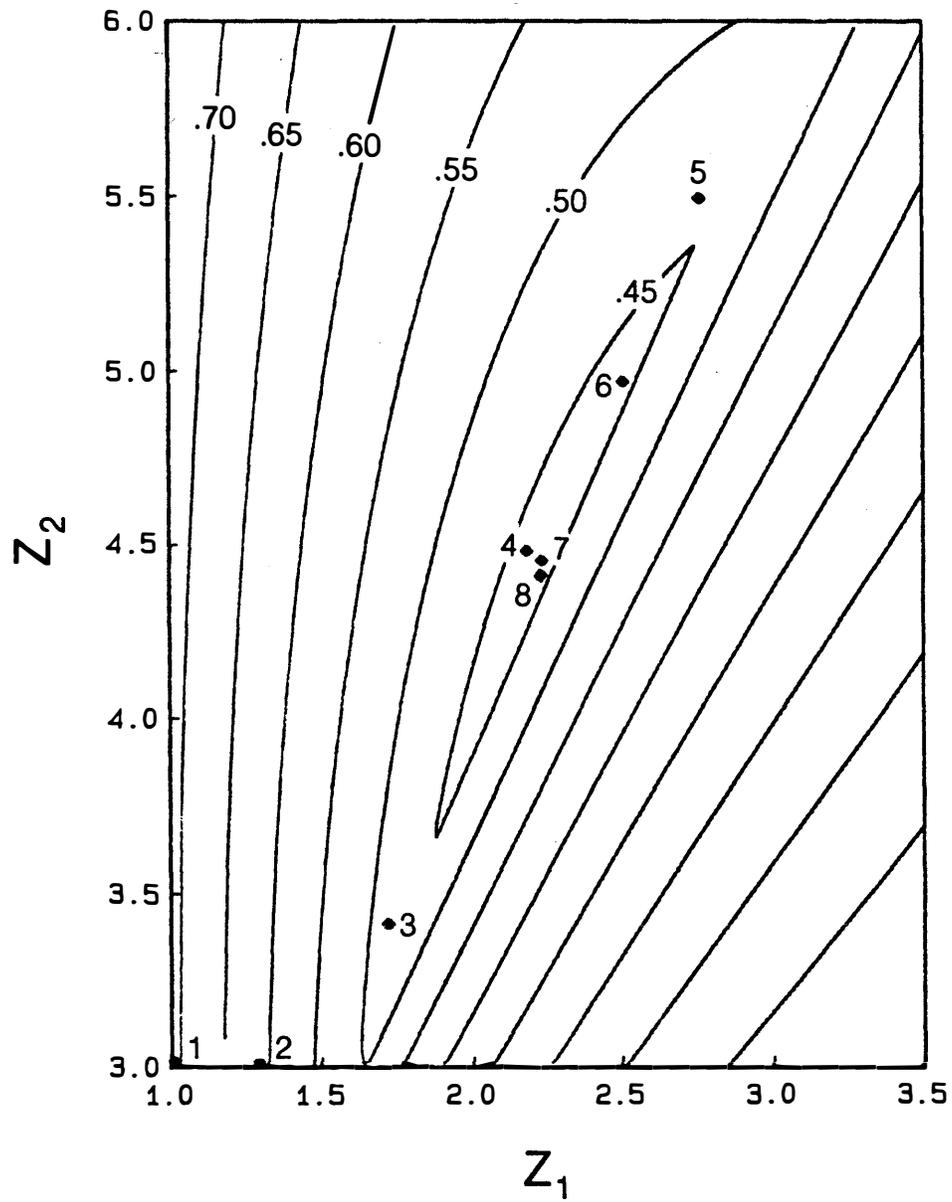


Fig. 7.2 Minimax contours for problem M1 (a two-dimensional singular minimax problem arising from optimization of the two-section transmission-line transformer). Eight iterations using exact gradients are illustrated.

optimization process and the solution given by Bandler, Kellermann and Madsen (1985) which used exact derivatives. If the gradients were estimated entirely by perturbations, 24 function evaluations would have to be performed. Using approximate gradients, the solution, shown in Fig. 7.3, required only 18 function evaluations.

#### Problem M2

For the same two-section transformer, we can define a regular minimax problem by choosing  $Z_1$  and  $l_1$  as variables while keeping  $Z_2$  and  $l_2$  at their optimal values (Bandler, Kellermann and Madsen 1985). Figs. 7.4 and 7.5 illustrate, respectively, the solution using exact derivatives and the solution using approximate gradients.

#### Problems M3 and M4

Two examples are considered of the design of multi-coupled cavity filters. These filters have been discussed in general in Chapter 4. Example M3 is a 4th-order filter having 4 designable couplings. M4 is of 6th-order and has 6 variables. The reflection coefficient in the passband is minimized and the transducer loss over the stopband is maximized.

#### Problems M5, M6 and M7

This is a test problem proposed by Brent (1972) for which the Newton-Raphson method is not globally convergent. We wish to solve the system of equations

$$\begin{aligned} h_1 &= 4(x_1 + x_2) = 0, \\ h_2 &= (x_1 - x_2)((x_1 - 2)^2 + x_2^2) + 3x_1 + 5x_2 = 0. \end{aligned} \tag{7.13}$$

We treat  $h_1$  as a linear equality constraint which can be

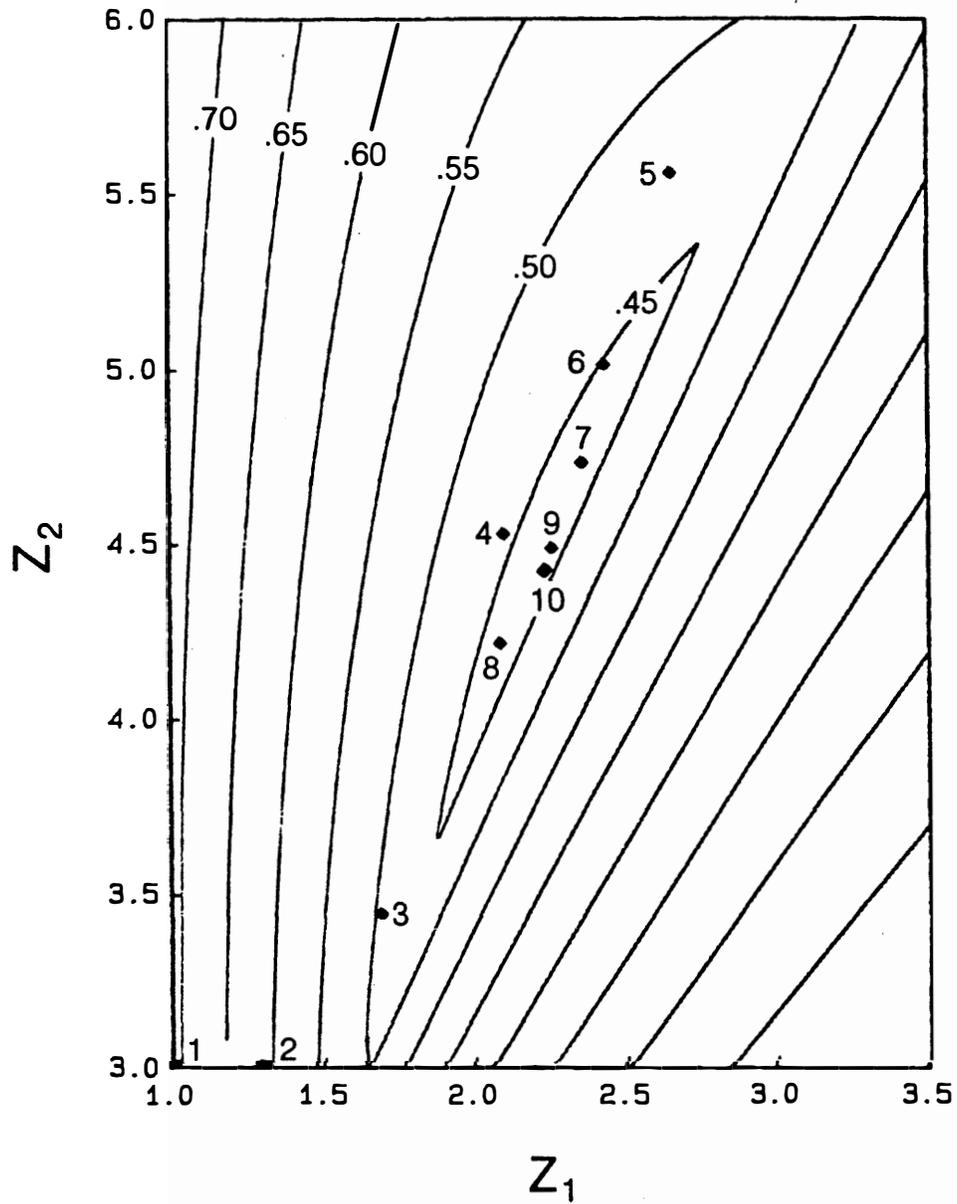


Fig. 7.3 Problem M1 is solved after 10 iterations of the minimax algorithm using approximate gradients.

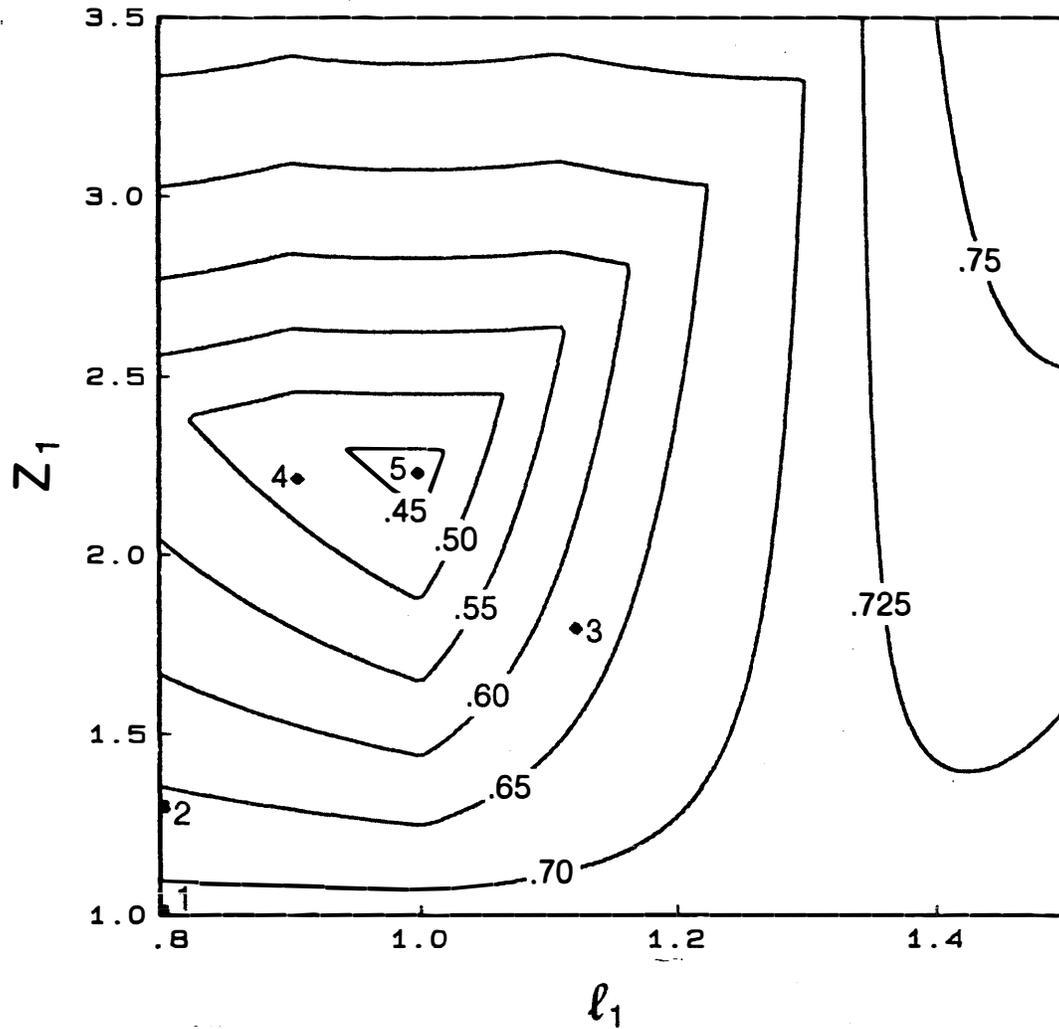


Fig. 7.4 Minimax contours for problem M2 (a two-dimensional regular minimax problem). Using exact gradients, a total of 8 iterations is required to reach the solution. The first 5 iterations are shown.

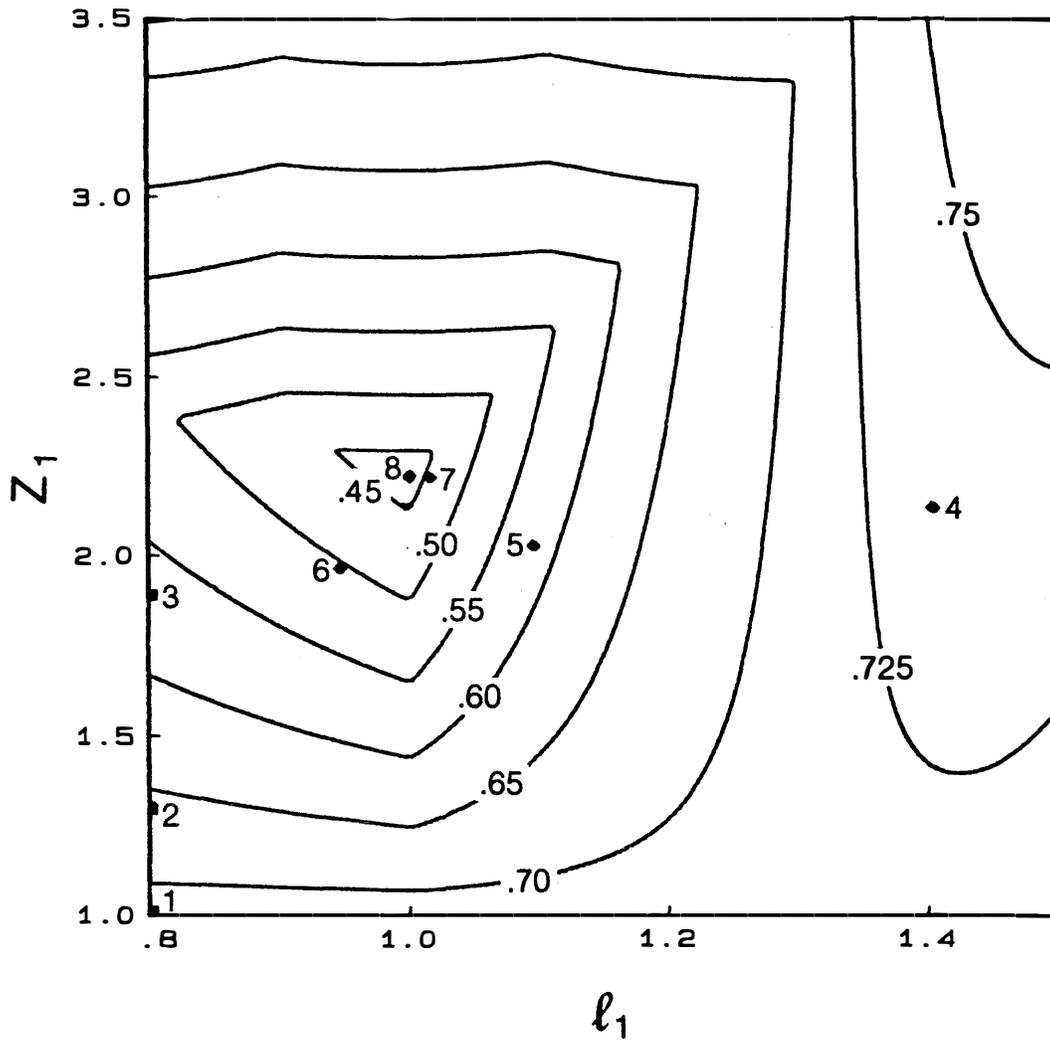


Fig. 7.5 Problem M2 is solved after 12 iterations of the minimax algorithm using approximate gradients. The first 8 iterations are illustrated.

handled directly by the minimax algorithm. We also define two error functions as  $f_1 = h_2$ ,  $f_2 = -h_2$ .

The problems M5, M6 and M7 correspond to three starting points used to solve (7.13), namely  $[2 \ 2]^T$ ,  $[2 \ 0]^T$  and  $[2 \ 1]^T$ , respectively.

### 7.3.3 Worst-Case Design of a Microwave Amplifier

The worst-case fixed tolerance design of a microwave amplifier is considered. Section 5.3.1 has addressed worst-case design in general. The amplifier, as shown in Fig. 7.6, consists of an NEC70000 FET and five transmission-lines. The FET is characterized by tabulated scattering parameters provided by the manufacturer (see TOUCHSTONE 1985). The design variables are the characteristic impedance  $Z$  and the lengths  $\ell_i$  of the transmission-lines. The design specifications are given by

$$7.05\text{dB} \leq 20\log|S_{21}| \leq 8.2\text{dB}, \quad \text{for } \omega_j = 6, 7, \dots, 18\text{GHz}.$$

Assuming a five percent tolerance associated with each length  $\ell_i$ , we seek an optimally centered design through a minimax optimization, as

$$\underset{\phi^0}{\text{minimize}} \max_j \max_k \{f_j(\phi^k)\} \quad (7.14)$$

where the error functions  $f_j$ ,  $j = 1, 2, \dots, 26$ , are derived from the upper and lower specifications at 13 frequency points. The vertices of the tolerance region are considered as candidates for the worst cases, denoted by  $\phi^k$ .

The worst-case design was accomplished by two phases of optimization. In the first one, we predicted an initial set of

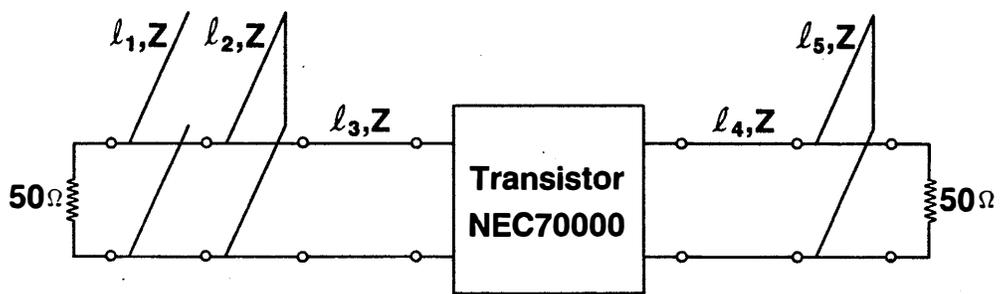


Fig. 7.6 A microwave amplifier.

worst-case vertices by first-order changes. For each  $f_j$ , a  $\phi^j$  was defined by

$$\phi_i^j = \phi_i^0 + \mu_i^j \epsilon_i, \quad \mu_i^j = \text{sign}(\partial f_j / \partial \phi_i), \quad i = 1, \dots, n, \quad (7.15)$$

where  $\epsilon_i$  is the tolerance associated with  $\phi_i$ . The derivatives at the starting point, which were also required for gradient approximation, were estimated by perturbations. Consequently, 26 worst cases (one for each  $f_j$ ) were considered and the minimax problem

$$\text{minimize}_{\phi^0} \max_j \{f_j(\phi^j)\} \quad (7.16)$$

was solved. At the solution, by using (7.15) with respect to the new nominal point, we found that 10 of the worst cases had changed (i.e., the signs of some  $\partial f_j / \partial \phi_i$  had changed). The new vertices were added to the worst-case set. The corresponding old vertices were kept, instead of replaced, in order to stabilize the algorithm. We had, therefore, a total of 36 worst cases. A second optimization was performed and at the solution the worst-case set was found to be complete (i.e., no more sign change in (7.15)).

The nominal parameter values at the starting point and the final solution are given in Table 7.2. The total number of function evaluations is 280, opposed to 585 required if perturbations were used throughout the optimization. Fig. 7.7 depicts the worst-case envelop at the solution.

#### 7.3.4 Design of a 5-Channel Multiplexer

Manifold multiplexer optimization has been discussed in Section 4.4, where a 16-channel contiguous band multiplexer was presented as an example.

TABLE 7.2  
PARAMETER VALUES OF THE MICROWAVE AMPLIFIER

---

Parameter	Starting Point	Solution
$l_1$	52.96	69.01
$l_2$	148.13	152.01
$l_3$	26.80	18.48
$l_4$	24.01	5.10
$l_5$	46.63	36.49
Z	81.27	126.39

---

The starting point is a minimax nominal design.

---

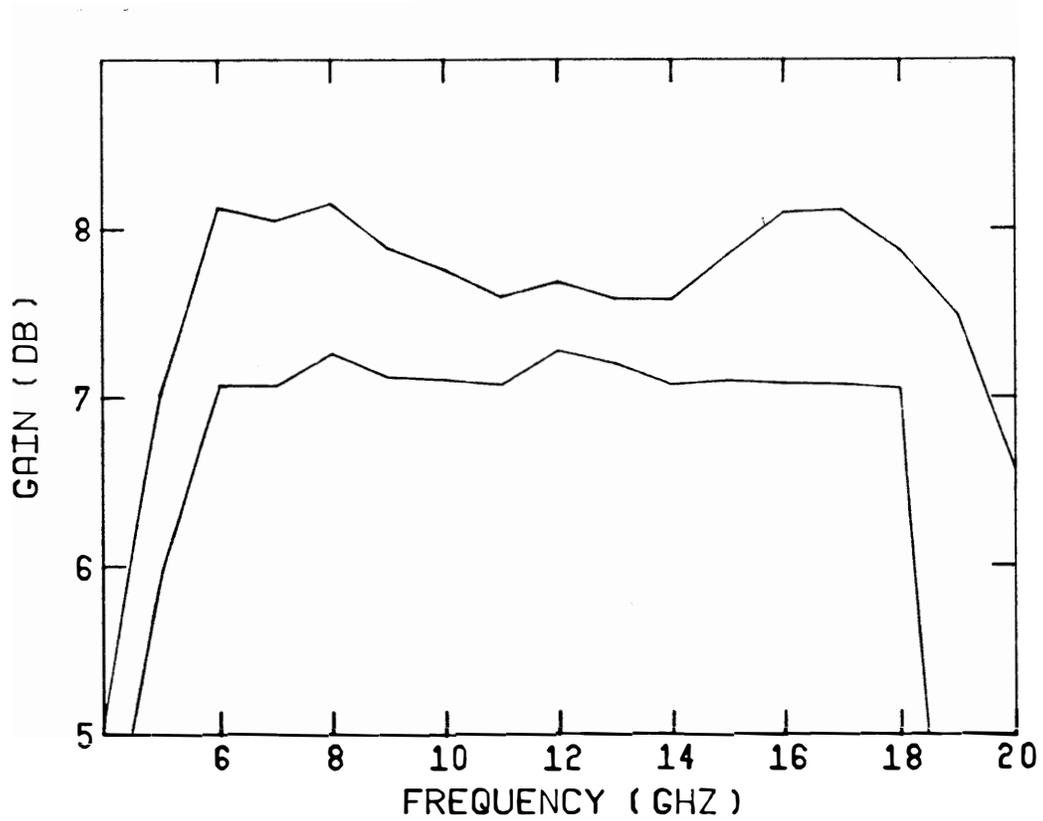


Fig. 7.7 Worst-case envelope for the amplifier response at the centered solution.

A minimax solution of a 5-channel 11GHz noncontiguous band multiplexer was given in detail by Bandler, Kellermann and Madsen (1985). To obtain the exact sensitivities required, the theory due to Bandler, Daijavad and Zhang (1986) was implemented in a computer program which has taken months of effort to develop and test. Furthermore, because the sensitivity expressions depend highly on the circuit structure and vary from component to component, every change to the problem, such as assigning different variables, requires expert modification to the software. In fact, sensitivities with respect to all possible variables were computed even though some of them have not been actually used, otherwise the coding scheme would have become unmanageable. Large amounts of computer memory were required to store various adjoint solutions and intermediate expressions.

By utilizing our gradient approximation, it is possible to efficiently design a multiplexer without all these troubles associated with computing the exact sensitivities. The complexity and size of the program can therefore be considerably reduced. It is obvious that to evaluate responses alone is more straightforward than to evaluate responses and sensitivities simultaneously.

The 5-channel multiplexer can be an excellent illustration of efficient gradient approximations for two reasons. First, it involves 75 variables and, therefore, to rely on perturbations would be prohibitively expensive. To be more specific, suppose that we use the initial parameter values and specifications suggested by Bandler, Kellermann and Madsen (1985). The multiplexer res-

ponses at the starting point are depicted in Fig. 7.8. An optimization after 50 iterations resulted in the responses of Fig. 7.9 (45 seconds on the FPS 264 via IBM 4381), when exact sensitivities were provided. To reach a similar result relying on perturbations for gradients, we would have to compute multiplexer responses 3800 times ( $50 \times 76$ ). We will show that efficient gradient approximations reduce the number of response evaluations significantly.

Also, this example is naturally suited for the use of the weighted Broyden update described in Section 7.2.3. From Fig. 7.9 it is intuitively obvious that the response functions at lower frequencies should be almost independent of the variables that are related to the filters of channels 1 and 2 (channel 1 has the highest center frequency). Similarly, the responses at higher frequencies are almost independent of the variables related to the filters of channels 3, 4 and 5. We will show that the use of appropriate weights improves the performance of the optimization.

The center frequencies and bandwidths of the five channels are given in Table 7.3. In the following experiments, all the channel filters start with the same 6th-order coupling matrix:

$$\mathbf{M} = \begin{bmatrix} 0 & 0.62575 & 0 & 0 & 0 & 0 \\ 0.62575 & 0 & 0.57615 & 0 & 0 & 0 \\ 0 & 0.57615 & 0 & 0.32348 & 0 & -0.74957 \\ 0 & 0 & 0.32348 & 0 & 1.04102 & 0 \\ 0 & 0 & 0 & 1.04102 & 0 & 1.04239 \\ 0 & 0 & -0.74957 & 0 & 1.04239 & 0 \end{bmatrix}.$$

The filters are lossy with an estimated Q factor of 12000. The

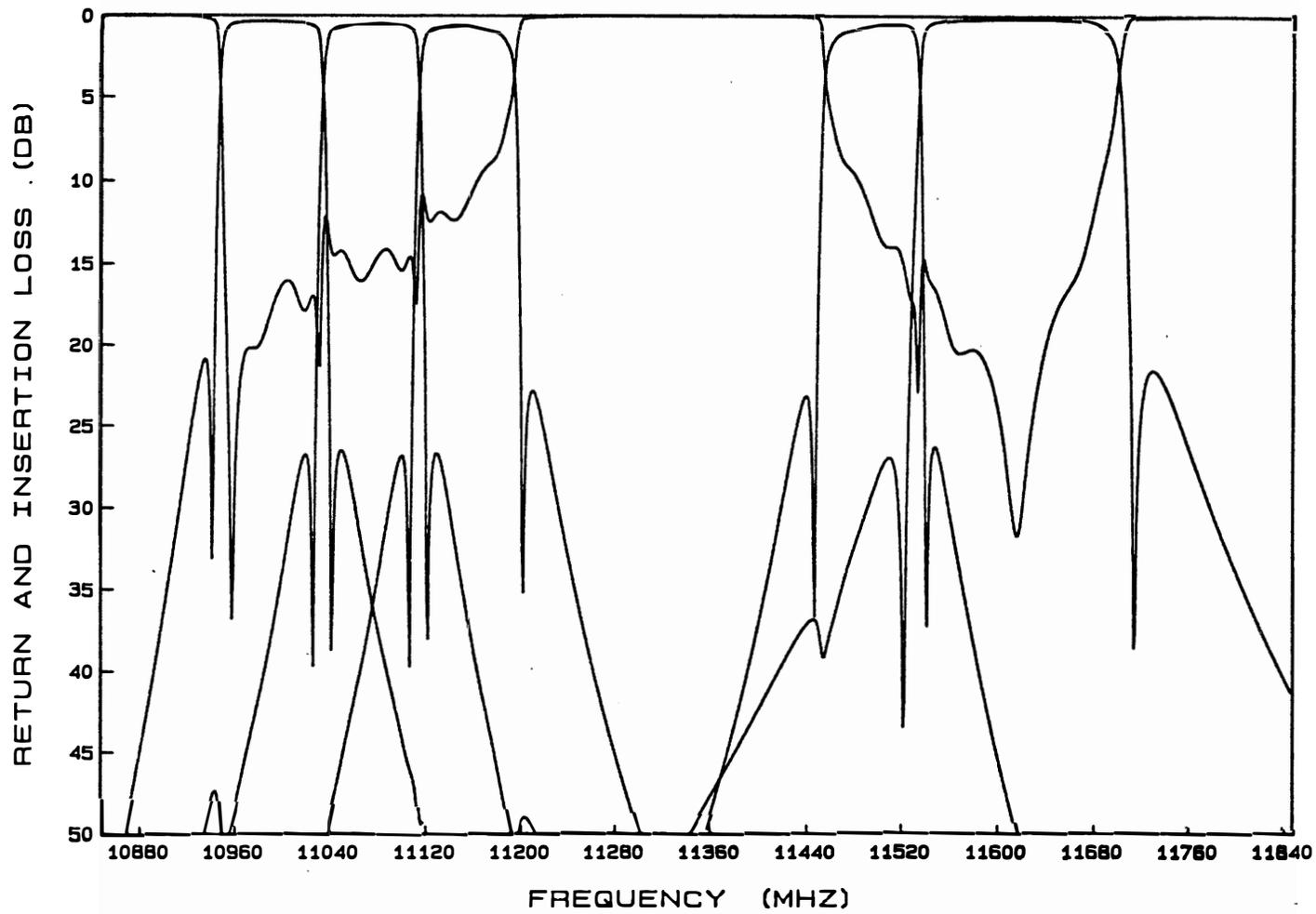


Fig. 7.8 Responses of the 5-channel, 11GHz multiplexer at the starting point.

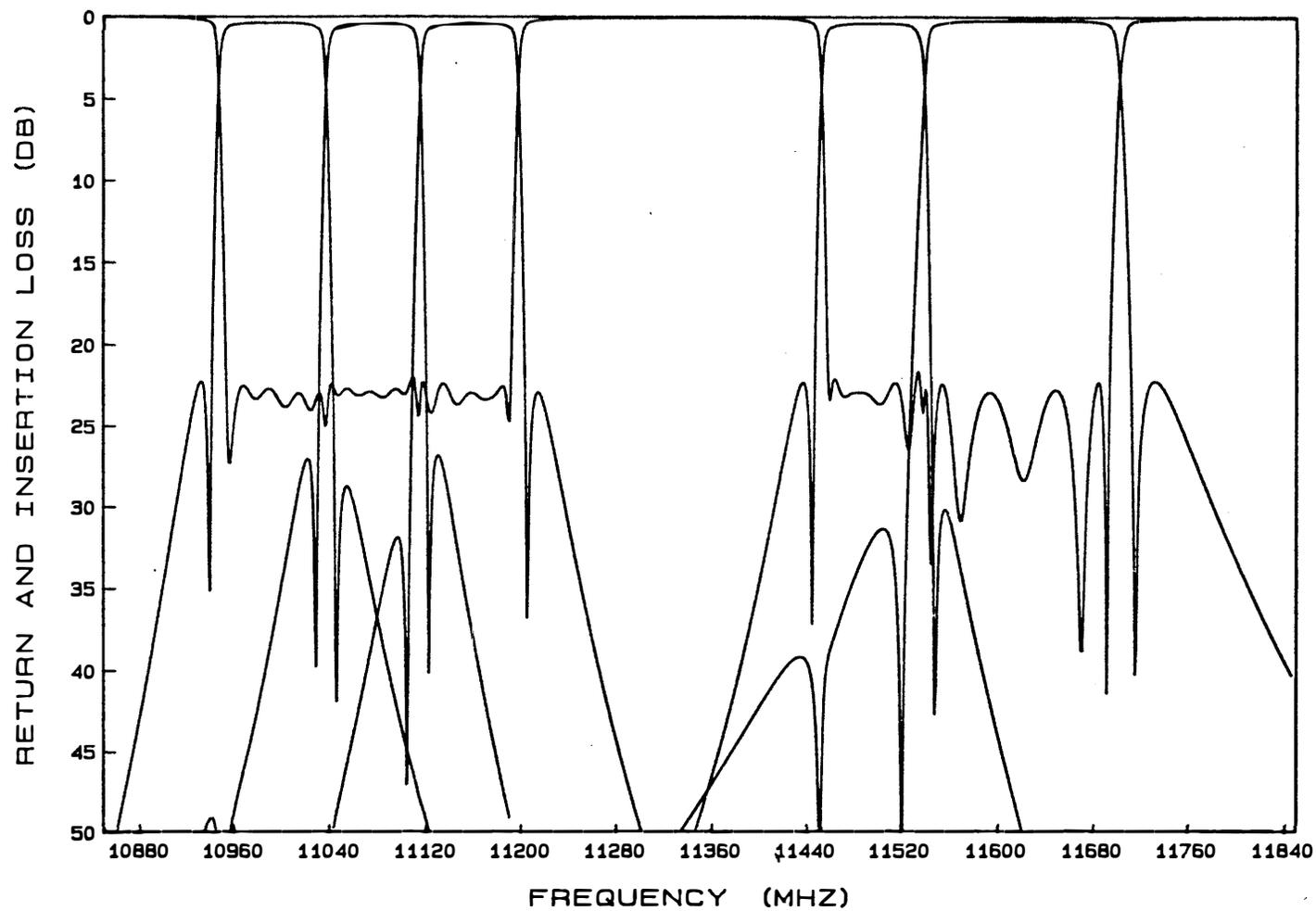


Fig. 7.9 Responses of the 5-channel multiplexer obtained after 50 optimization iterations using exact gradients. It would require 3800 multiplexer response evaluations to estimate the gradients entirely by perturbations.

TABLE 7.3  
MULTIPLEXER CENTER FREQUENCIES AND BANDWIDTHS

---

Channel	Center Frequency (MHz)	Bandwidth (MHz)
1	11618.5	154
2	11495	76
3	11155	76
4	11075	76
5	10992.5	81

---

initial spacing for the waveguide section associated with each channel is set equal to half the guide wavelength evaluated at the center frequency of the corresponding channel. The input and output transformer ratios start from  $n_1^2 = 0.68820$  and  $n_2^2 = 2.04417$ .

A lower specification of 20dB is imposed on the common port return loss, for which a total of 52 frequency points is used. These points are spaced 10MHz apart in the passband of each channel with additional single frequencies at the crossover of two contiguous channels. A lower specification of 20dB on the transition band insertion loss is also imposed at frequencies of 10935, 11210, 11215, 11440, 11442, 11712 and 11725MHz. Table 7.4 summarizes the parameter values of the solution obtained using exact gradients.

#### Experiment 1

In the first experiment, perturbations were used only at the starting point but not during the optimization. The approximation of gradients relied on the Broyden update with special iterations, which was similar to the methods of Madsen (1975) and Zuberek (1984). The optimization stopped after 266 response evaluations (81 seconds on the FPS 264), of which 75 were used for the initial perturbations. The responses at this solution as shown in Fig. 7.10 are obviously not as good as the ones shown in Fig. 7.9. The optimization has stopped prematurely. This experiment has demonstrated that the Broyden update may not be sufficient for large-scale and/or highly nonlinear problems.

#### Experiment 2

In a second experiment, regular corrections were provided

TABLE 7.4  
MULTIPLEXER PARAMETERS OPTIMIZED USING EXACT GRADIENTS

Parameter	Ch. 1	Ch. 2	Ch. 3	Ch. 4	Ch. 5
$M_{11}$	-0.0417	0.2194	-0.0859	0.0454	0.0459
$M_{22}$	-0.0708	0.0301	-0.0596	-0.0098	0.0310
$M_{33}$	-0.0209	-0.0215	-0.0119	-0.0097	0.0069
$M_{44}$	-0.0196	-0.0621	0.0158	-0.0121	-0.0070
$M_{55}$	0.0414	-0.0172	0.0121	0.0023	0.0141
$M_{66}$	0.0402	0.0117	0.0339	-0.0058	0.0104
$M_{12}$	0.7598	0.7383	0.7091	0.6115	0.6592
$M_{23}$	0.5723	0.6096	0.5845	0.5551	0.5873
$M_{34}$	0.4239	0.4221	0.4086	0.3048	0.3644
$M_{36}$	-0.5326	-0.6346	-0.6021	-0.7519	-0.6948
$M_{45}$	0.8971	1.0266	0.9916	1.0317	1.0468
$M_{56}$	1.1023	1.1518	1.1715	1.0558	1.1186
$n_1$	1.0547	0.9358	0.9343	0.8188	0.8031
$n_2$	1.4350	1.4311	1.4286	1.4153	1.4120
$\rho$	0.7033	0.6039	0.9219	0.7191	0.7295

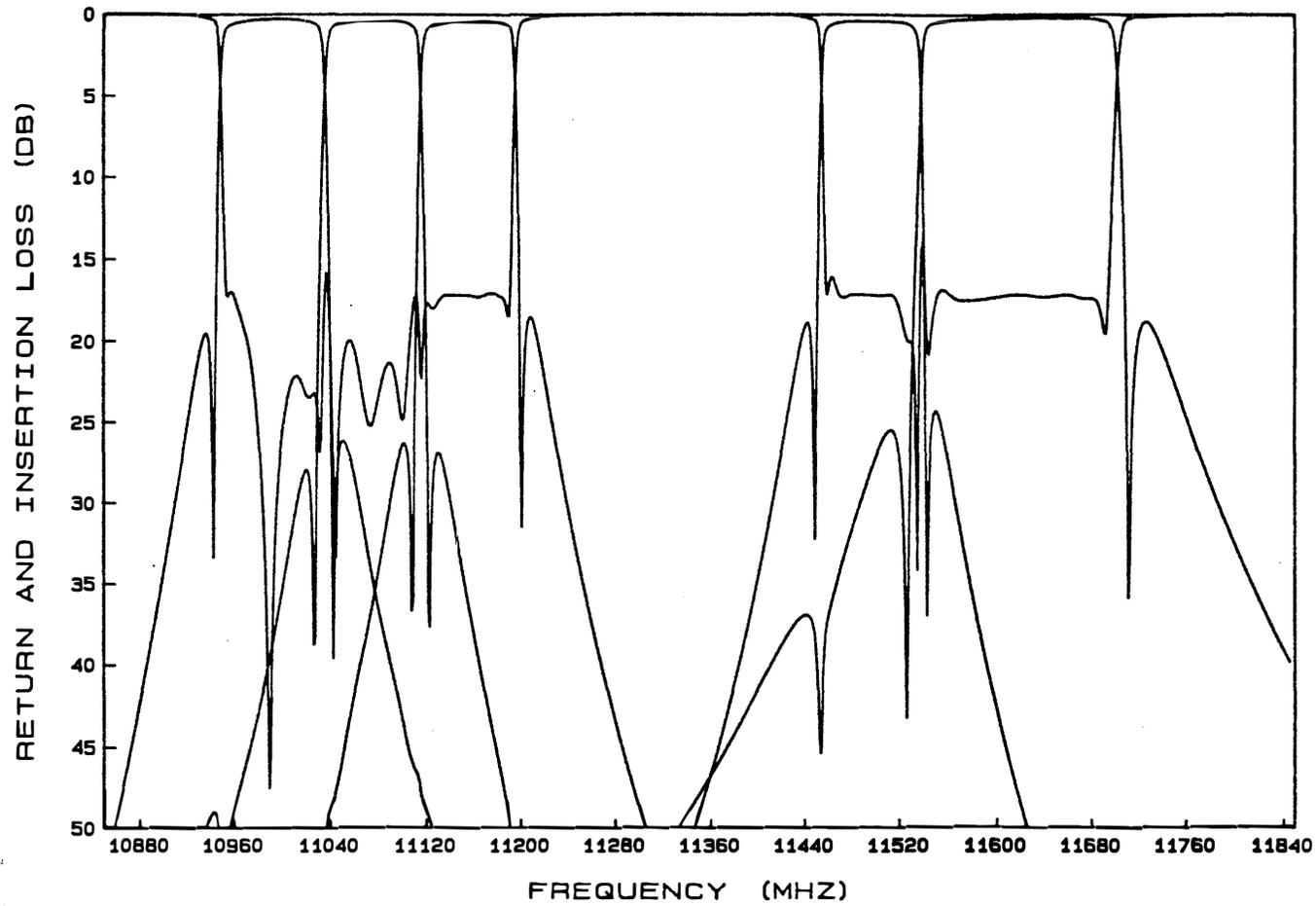


Fig. 7.10 Responses of the 5-channel multiplexer obtained using only the Broyden update and special iterations for gradient approximation. The optimization has stopped prematurely.

by perturbations for every 20 iterations. After 500 response evaluations, of which 375 were used for perturbations, we obtained the responses shown in Fig. 7.11. Continuing the process for another 500 response evaluations the responses shown in Fig. 7.12 were achieved, which are as good as the ones in Fig. 7.9. Table 7.5 summarizes the parameter values of the final solution. From the starting point, a total of 1000 response evaluations (298 seconds on the FPS 264) was performed. Recall that 3800 response evaluations would be required if the gradient calculations were simply replaced by perturbations.

### Experiment 3

Experiments one and two have both used the original Broyden update. Our third experiment demonstrates the weighted update described in Section 7.2.3. For this formula, a weight  $w_{j_i}$  is set to zero when we know that a function  $f_j$  is almost independent of a variable  $x_i$ . For instance, the insertion loss of channels 3, 4 and 5 and the common port return loss over the passbands of these channels are almost independent of the filter couplings in channels 1 and 2. Similarly, the responses within the frequencies of channels 1 and 2 are almost independent of the filter couplings in channels 3, 4 and 5. Therefore, we set the corresponding weights to zero.

Utilizing the weighted update, we optimize the multiplexer without any regular correction by perturbations. The responses shown in Fig. 7.13 were obtained after 500 response evaluations (166 CPU seconds on the FPS 264). By comparing this result with

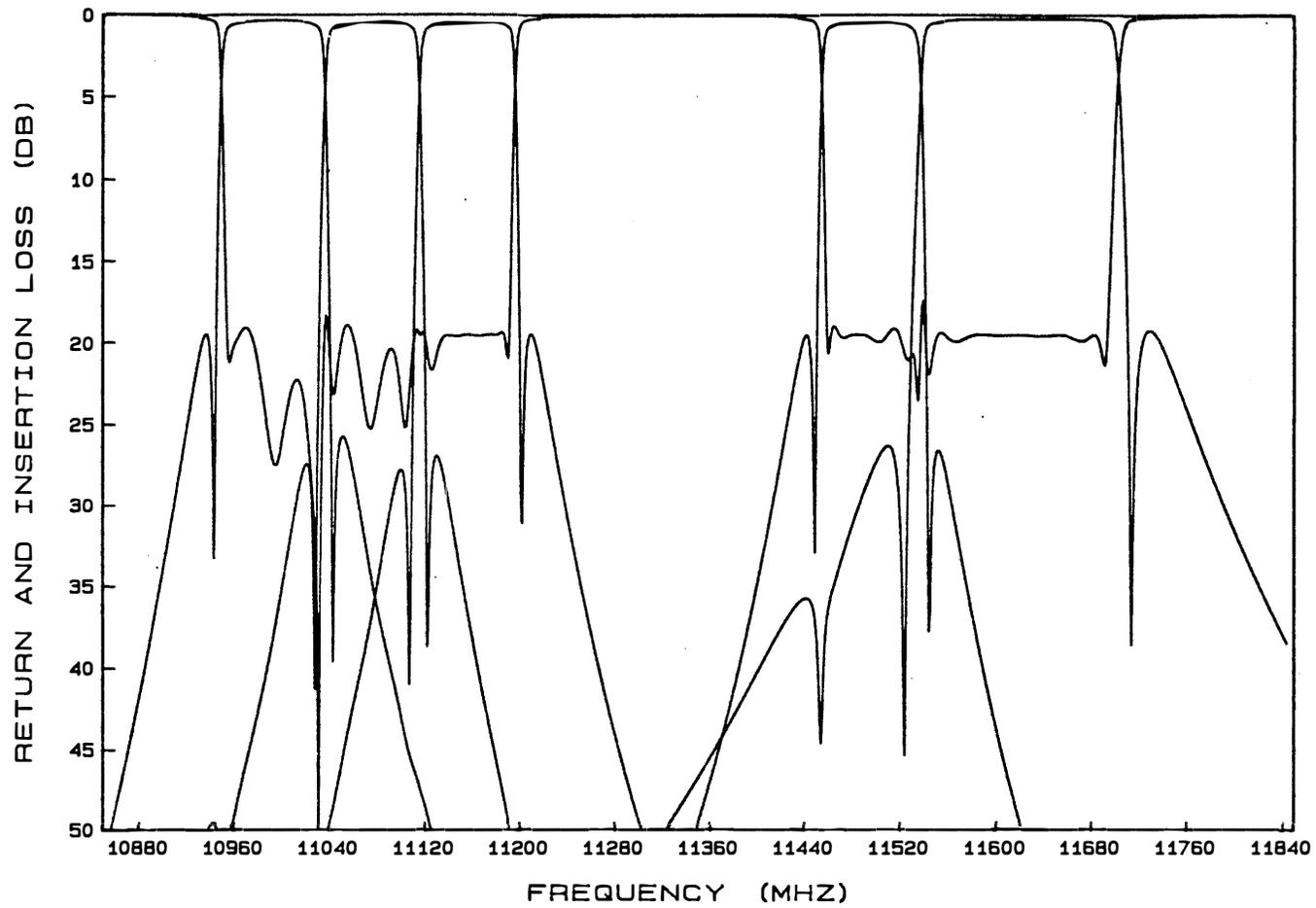


Fig. 7.11 Responses of the 5-channel multiplexer obtained after 500 response evaluations. Approximate gradients were used with regular corrections provided by perturbations.

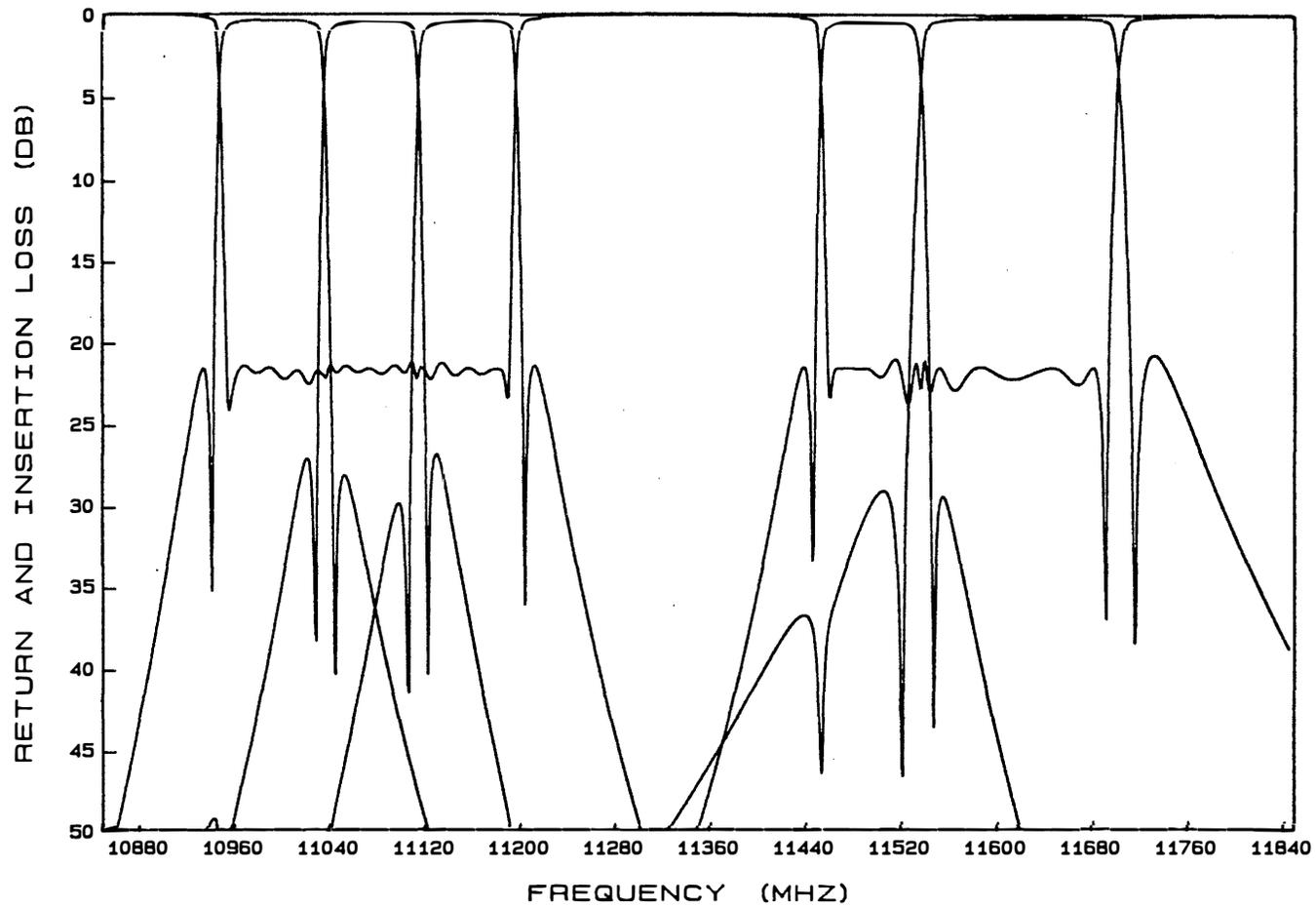


Fig. 7.12 Responses of the 5-channel multiplexer at the solution of experiment 2, after 500 response evaluations from the point shown in Fig. 7.11 (1000 response evaluations from the start).

TABLE 7.5

MULTIPLEXER PARAMETERS OPTIMIZED USING APPROXIMATE GRADIENTS

Parameter	Ch. 1	Ch. 2	Ch. 3	Ch. 4	Ch. 5
$M_{11}$	-0.0432	0.1801	-0.0788	0.0384	0.0401
$M_{22}$	-0.0526	0.0294	-0.0514	-0.0067	0.0254
$M_{33}$	-0.0082	-0.0178	-0.0082	-0.0055	0.0044
$M_{44}$	0.0158	-0.0582	0.0160	-0.0064	-0.0102
$M_{55}$	0.0160	-0.0206	0.0090	0.0021	0.0038
$M_{66}$	-0.0255	0.0100	-0.0248	-0.0037	0.0066
$M_{12}$	0.7427	0.7077	0.6969	0.6124	0.6495
$M_{23}$	0.5796	0.5951	0.5815	0.5567	0.5800
$M_{34}$	0.3855	0.3876	0.3780	0.3050	0.3491
$M_{36}$	-0.6314	-0.6699	-0.6540	-0.7520	-0.7083
$M_{45}$	0.9657	1.0338	1.0064	1.0325	1.0324
$M_{56}$	1.1330	1.1279	1.1346	1.0553	1.0983
$n_1$	0.9976	0.8915	0.8997	0.7993	0.7862
$n_2$	1.4416	1.4236	1.4223	1.4140	1.4154
$\ell$	0.6958	0.6132	0.9235	0.7228	0.7360

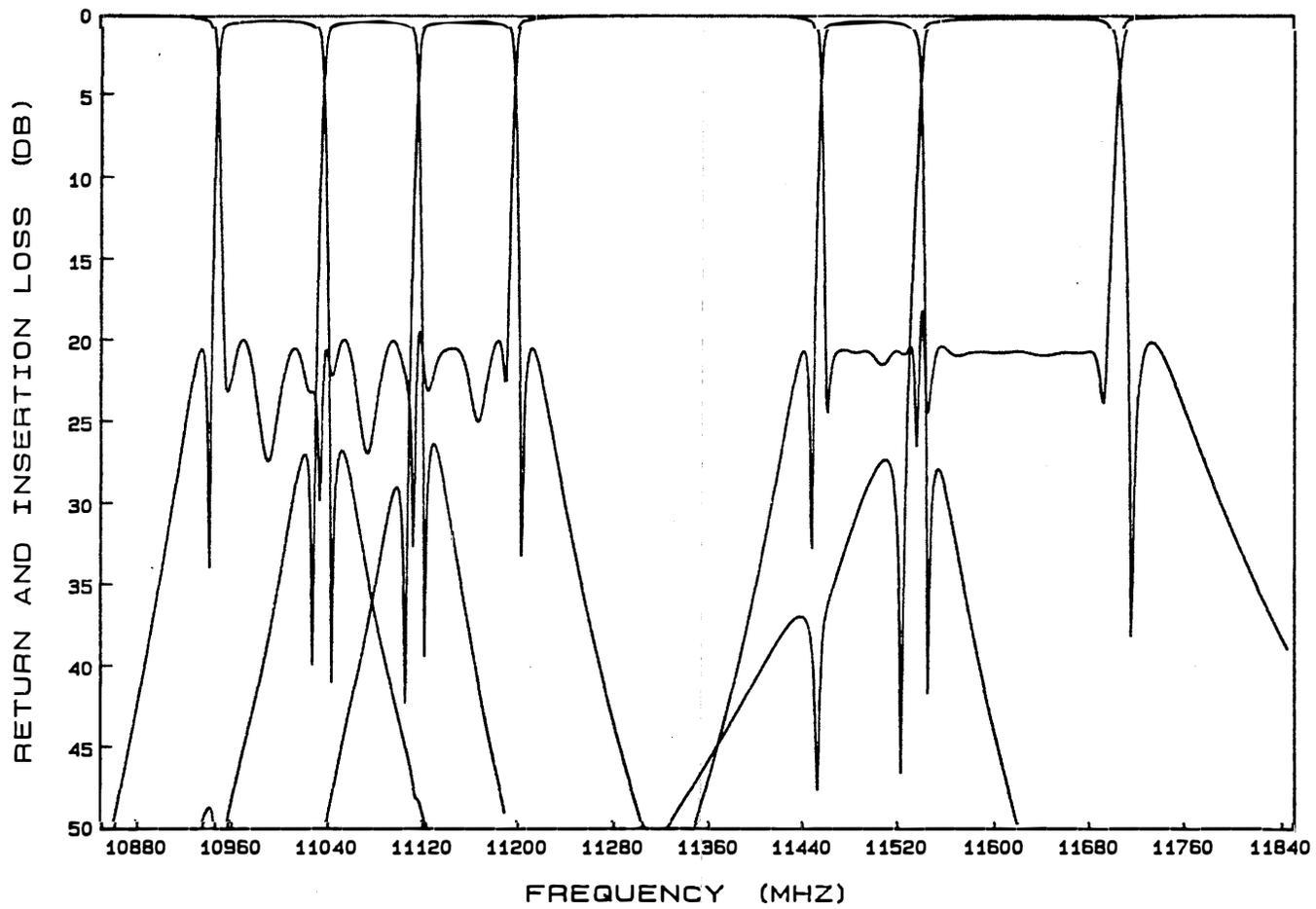


Fig. 7.13 Responses of the 5-channel multiplexer at the solution of experiment 3. Utilizing the weighted update, it required only 500 response evaluations.

experiment 1 we can clearly see that the use of appropriate weights has prevented the optimization from stopping prematurely. We can also conclude from a comparison between experiments 2 and 3 (also, between Figs. 7.11 and 7.13) that applying the weighted update has effectively reduced the use of time-consuming perturbations.

#### 7.4 AN $\ell_1$ ALGORITHM USING APPROXIMATE GRADIENTS

##### 7.4.1 Implementation of a Fortran Subroutine

We have also integrated gradient approximations with the Hald and Madsen (1985) algorithm for the  $\ell_1$  problem as

$$\underset{\mathbf{x}}{\text{minimize}} \sum_{j=1}^m |f_j(\mathbf{x})| \quad (7.17)$$

subject to possible linear constraints. Similar to the minimax algorithm in its structure, the  $\ell_1$  algorithm also consists of two stages, Stage 1 being a trust region Gauss-Newton method which provides global convergence and Stage 2 a quasi-Newton method which is intended to achieve a fast rate of convergence near a solution.

The original algorithm requires a user-supplied subroutine which calculates the values of the error functions as well as their first-order derivatives. This is replaced in the new algorithm by an interface which organizes gradient approximations as outlined in Section 7.2.6. From the user's point of view, the  $\ell_1$  package is almost identical to the minimax package which has been described in Section 7.3.1. An external subroutine is required to supply the function values. The user specifies, via common block parameters, the method of obtaining the initial approximate Jacobian and the

use of regular perturbations during optimization. Different rates of perturbations for Stage 1 and Stage 2 are allowed.

The KMOS software library described in Chapter 2 has also included the Fortran implementation of this algorithm.

#### 7.4.2 Performance on Some Test Problems

We have tested the  $\ell_1$  algorithm using approximate gradients on a large variety of problems. Some of the tests are described in this section. Convergence was achieved for all these problems and our solutions agree with the exact solutions to at least five significant figures. The computational effort required by the new algorithm and the effort required for estimating gradients entirely by perturbations are compared in Tables 7.6 and 7.7.

##### Problem L1

The two-section transmission-line transformer, which we have used in Problem M1 of Section 7.3.2, is considered here for parameter identification using the  $\ell_1$  optimization. The reflection coefficient of the transformer at the solution of Problem M1 was taken as the measurement, from which we attempt to identify the values of  $Z_1$  and  $Z_2$ . Figs. 7.14 and 7.15 illustrate, on the  $\ell_1$  contours, two solutions, one obtained with the gradient estimated entirely by perturbations and the other by our approximate gradient algorithm.

##### Problem L2

This is a data-fitting problem considered by Madsen (1975)

TABLE 7.6  
 COMPARISON OF COMPUTATIONAL EFFORT FOR EXAMPLES L1 TO L4

Problem	Number of Function Evaluations	
	Entirely by Perturbations	By the New Algorithm
L1	42 (14)	27 (19)
L2	54 (9)	32 (19)
L3	105 (15)	63 (40)
L4	71 (17)	65 (48)

The entries in parentheses are numbers of optimization iterations.

TABLE 7.7  
COMPARISON OF COMPUTATIONAL EFFORT FOR EXAMPLE L5

Size of the Problem	Number of Function Evaluations		
	Case 1	Case 2	Case 3
n = 5	36 (6)	17 (9)	13 (7)
n = 10	66 (6)	25 (10)	19 (7)
n = 20	126 (6)	39 (13)	29 (7)

Case 1: The gradients were estimated entirely by perturbations.

Case 2: Using the Broyden update without weights.

Case 3: Using the weighted update.

The entries in parentheses are numbers of optimization iterations.

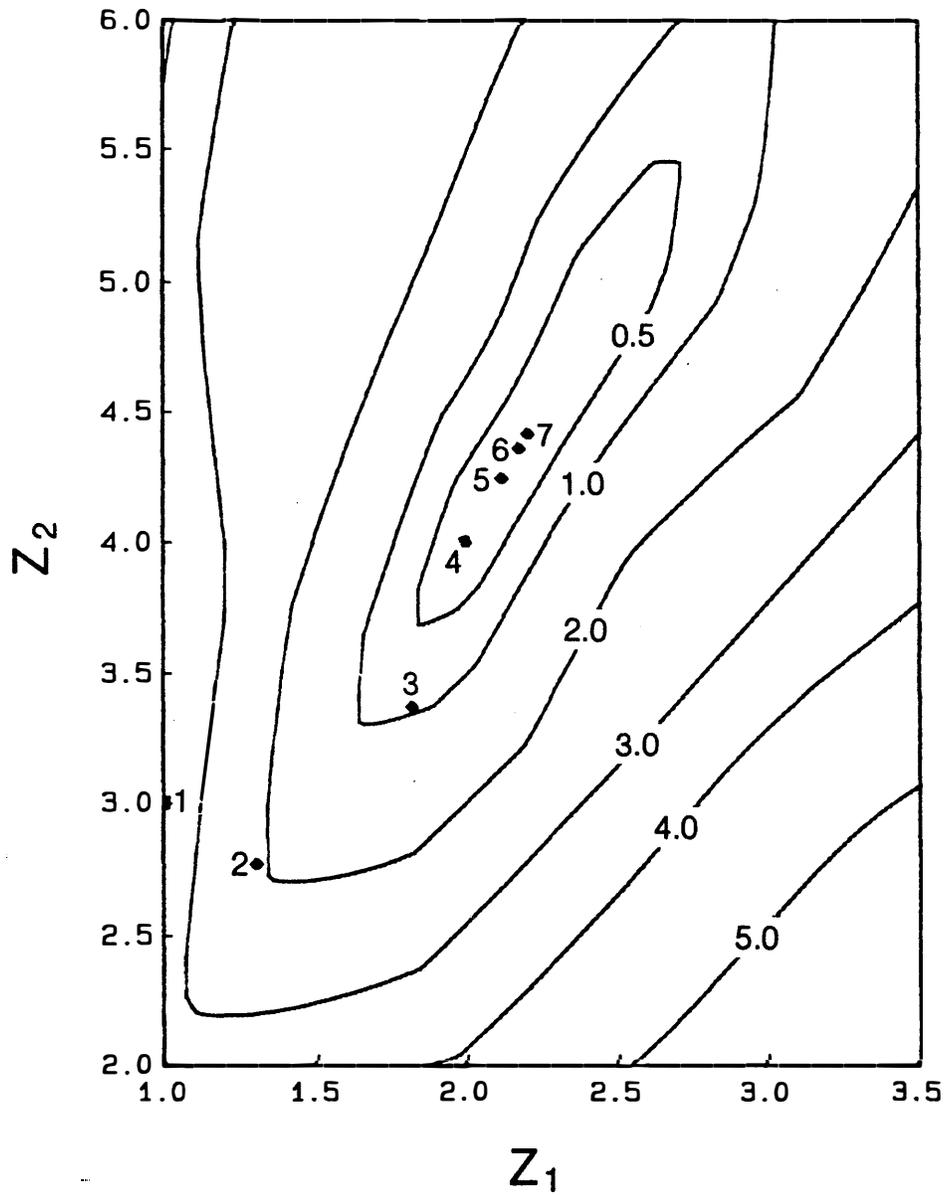


Fig. 7.14  $\lambda_1$  contours for problem L1 arising from parameter identification of the two-section transmission-line transformer. Using perturbations for the gradients, the solution required 14 iterations (42 function evaluations). The first 7 iterations are illustrated.

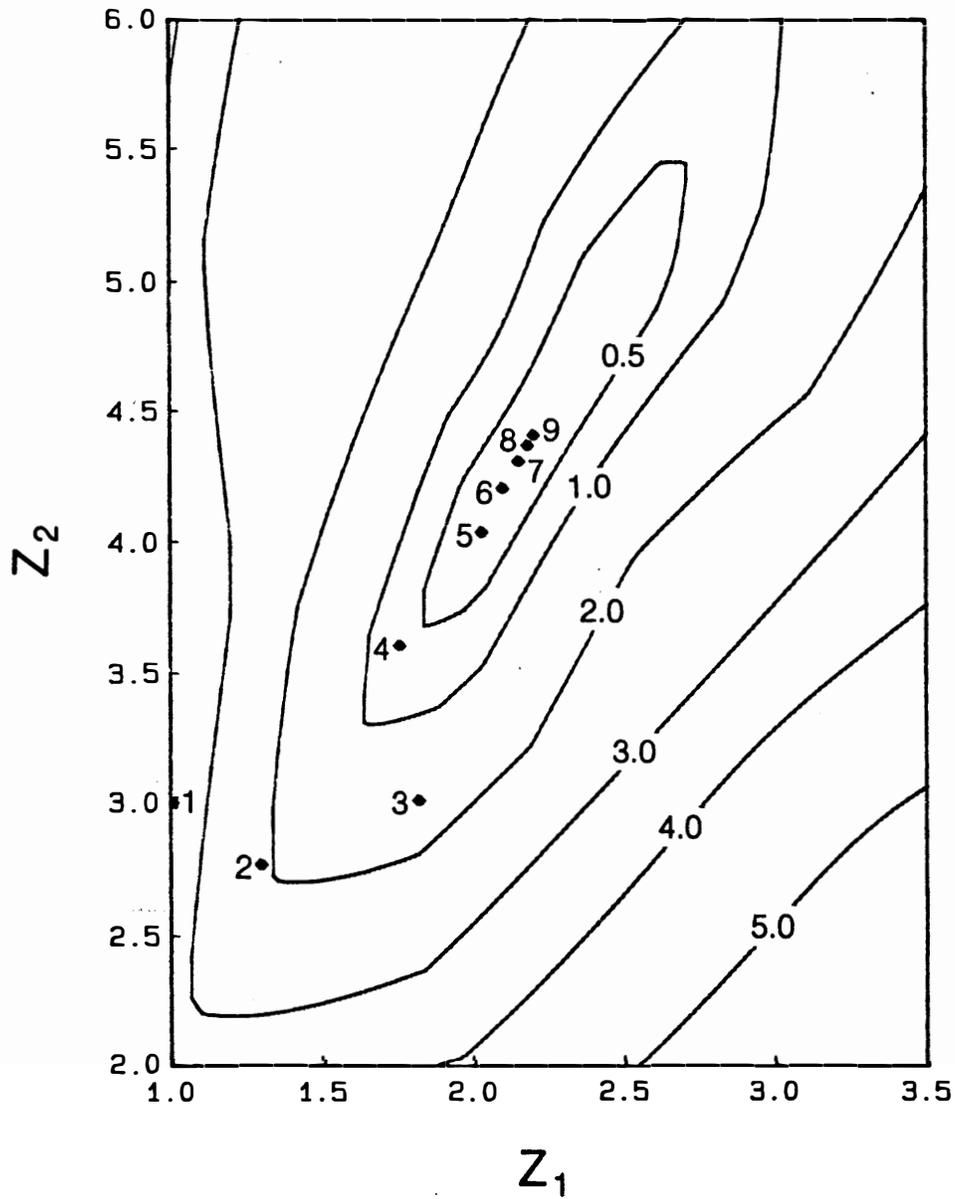


Fig. 7.15 Problem L1 is solved after 19 iterations (27 function evaluations) of the  $\ell_1$  algorithm using approximate gradients. The first 9 iterations are illustrated.

in which  $e^y$  is approximated by a third-order rational function over the interval  $-1 \leq y \leq 1$ . The error functions are

$$f_j(\mathbf{x}) = \frac{x_1 + x_2 y_j}{1 + x_3 y_j + x_4 y_j^2 + x_5 y_j^3} - \exp(y_j),$$

$$y_j = -1 + 0.1(j - 1), \quad j = 1, \dots, 21,$$
(7.18)

where  $\exp(y)$  represents  $e^y$ .

#### Problem L3

This is a problem due to El-Attar, Vidyasagar and Dutta (1979) of finding a third-order model for a seventh-order system. The problem involves 6 variables and 51 functions.

$$f_j(\mathbf{x}) = x_1 \exp(-x_2 t_j) \cos(x_3 t_j + x_4) + x_5 \exp(-x_6 t_j) - y_j,$$

$$y_j = 0.5 \exp(-t_j) - \exp(-2t_j) + 0.5 \exp(-3t_j)$$

$$+ 1.5 \exp(-1.5t_j) \sin(7t_j) + \exp(-2.5t_j) \sin(5t_j),$$
(7.19)

$$t_j = 0.1(j - 1), \quad j = 1, \dots, 51.$$

This problem has also been solved by Bandler, Kellermann and Madsen (1987) using exact derivatives.

#### Problem L4

This problem involves a set of nonlinear equations given by

$$f_1(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 - 1,$$

$$f_2(\mathbf{x}) = x_1^2 + x_2^2 + (x_3 - 2)^2,$$

$$f_3(\mathbf{x}) = x_1 + x_2 + x_3 - 1,$$

$$f_4(\mathbf{x}) = x_1 + x_2 - x_3 - 1,$$

$$f_5(\mathbf{x}) = 2x_1^3 + 6x_2^2 + 2(5x_3 - x_1 + 1)^2,$$

$$f_6(\mathbf{x}) = x_1^2 - 9x_3.$$
(7.20)

El-Attar, Vidyasagar and Dutta (1979) have attempted an  $\ell_1$  solution to (7.20). Bandler, Kellermann and Madsen (1987) have

also considered this problem, and their solution was reported to be singular.

Problem L5

In Section 7.2.3, we have proposed a weighted Broyden update. Using the weighted update, possible special structure of a system, such as a sparse Jacobian, can be exploited. Consider a class of equations (Broyden 1965) given by

$$\begin{aligned} f_1(\mathbf{x}) &= (3 - 0.5x_1)x_1 + 2x_2 - 1, \\ f_j(\mathbf{x}) &= x_{j-1} - (3 - 0.5x_j)x_j + 2x_{j+1} - 1, \\ & \quad j = 2, 3, \dots, n-1, \\ f_n(\mathbf{x}) &= x_{n-1} - (3 - 0.5x_n)x_n - 1. \end{aligned} \tag{7.21}$$

In this tridiagonal system  $f_j$  is linear in  $x_i$  for all  $i \neq j$ . Following the discussion in Section 7.2.3, we define a set of weights as  $w_{j,j} = 1$  and  $w_{j,i} = 0$  for  $i \neq j$ . Using the weighted update we have solved (7.21) for  $n = 5, 10$  and  $20$ . The results in Table 7.7 show clearly that the weighted update is more efficient than the original Broyden formula in this example. The saving in computation becomes more significant as the size of the system increases. The potential advantages of the weighted update in practical  $\ell_1$  optimization will be further demonstrated later in this chapter through applications to FET modeling.

#### 7.4.3 Fault Location of a Mesh Network

Fault location of a resistive mesh network has been solved by Bandler, Kellermann and Madsen (1987) using an  $\ell_1$  algorithm which requires exact derivatives. As shown in Fig. 7.16, the

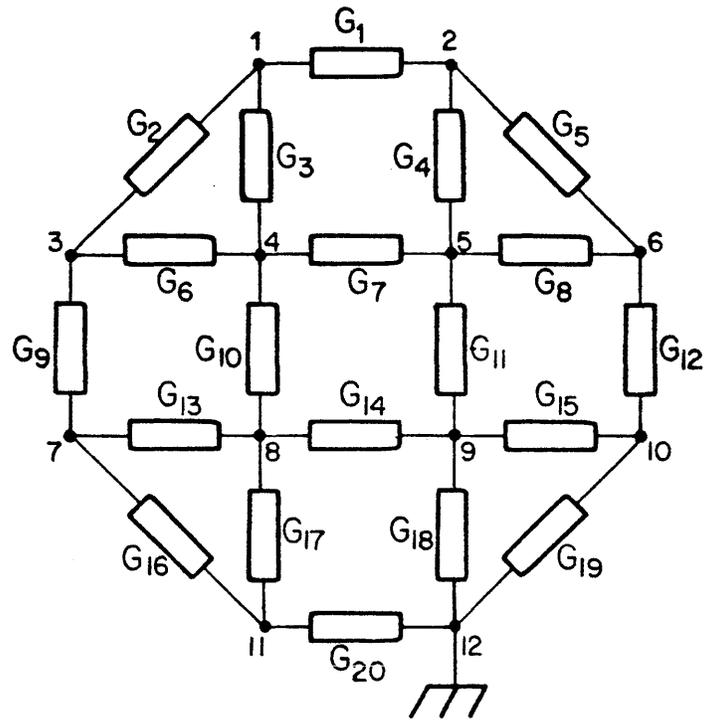


Fig. 7.16 A resistive mesh network.

network consists of 20 elements with the nominal values  $G_i = 1.0$  for  $i = 1, \dots, 20$ . Two faults are assumed in the network, namely  $G_2 = G_{18} = 0.5$ . A five percent tolerance is associated with each of the non-faulty elements. All outside nodes are assumed to be accessible and a single excitation is applied to node 1.

Utilizing our  $\ell_1$  algorithm with integrated gradient approximations, the faulty elements were correctly located after 34 network simulations. The same problem was also solved using perturbations for the gradients, which required a total of 147 network simulations.

#### 7.4.4 Multi-Circuit Modeling of a FET Device

In Chapter 6 we have described a novel approach to device modeling which exploits the unique properties of the  $\ell_1$  optimization and employs the concept of simultaneous processing of multiple circuits. Its application to the modeling of an actual FET device has also been presented in Section 6.4.2.

In the context of this chapter, the same FET modeling problem is solved, this time without calculating the exact sensitivities, to illustrate practical  $\ell_1$  optimization with integrated gradient approximations.

A detailed description of the small-signal equivalent circuit, the model parameters and the measurements of the FET device was given in Section 6.4.2. From 11 model parameters

$$\{R_g, R_d, L_s, \tau, R_{ds}, R_i, R_s, C_{gs}, C_{dg}, C_{ds}, g_m\}$$

the first four were selected as common variables. Three sets of

measurements on scattering parameters were utilized. The overall optimization problem involved 25 variables and 408 error functions.

In Section 6.4.2, we have solved the problem using exact sensitivities. The programming was quite involved because a comprehensive coding scheme was needed to identify the appropriate sensitivity expressions for the functions and variables associated with different circuits. It would be difficult to change the circuit topology or the variable designation without considerable labor. In comparison, a subroutine which calculates the function values only is much less complex.

Three experiments were conducted which have used different schemes to estimate the gradients. From the starting point given in Table 6.1, they have reached practically the same solution, which has also been given in Table 6.1.

In the first experiment the gradients were estimated solely by perturbations. A total of 468 circuit simulations were required to reach the solution.

In the second case, the Broyden update without weights was used. Regular corrections were also provided by perturbations for every five iterations. Only 128 circuit simulations were needed for this solution.

For the third experiment, we took advantage of an inherent decomposition in the multi-circuit formulation. The responses (and error functions) of one circuit are absolutely unrelated to the independent parameters of any other circuits. Obviously, the derivatives corresponding to such decoupled functions and variables are

always equal to zero. However, when we use the Broyden update without weights, these derivatives may be changed to some nonzero values, thus introducing apparent errors to the approximation. We can avoid this by using the weighted update. By assigning zero weights to decoupled functions and variables, we can keep the zero derivatives undisturbed throughout the optimization process. The application of this concept has reduced the use of perturbations and led to the solution after only 79 circuit simulations. This represents less than 1/5 of the simulations required by the first experiment as well as a 38% saving in computational effort as compared to the second experiment.

#### 7.5 CONCLUDING REMARKS

In this chapter, we have described a new approach to gradient approximations. Combining perturbations, the Broyden update and the special iterations, the new approach has significantly improved the computational efficiency as compared with the more conventional methods. A weighted update has also been proposed which exploits possible sparsity and decoupled structures of a system to further reduce the computations involved in estimating gradients. Integration of our approach with powerful gradient-based optimization techniques has been described and illustrated by the minimax and  $\ell_1$  implementations. The effectiveness and efficiency of our algorithm have been demonstrated through a large variety of problems. Examples of significant practical interest have been given in detail, including worst-case design, multiplexer

optimization, fault location and multi-circuit modeling.

Knowing that many CAD packages currently used in industry are not capable of providing exact sensitivities, the author strongly believes that efficient gradient approximations will contribute greatly to extending the application of advanced optimization tools to a much broader range of practical problems.

## CHAPTER 8

### CONCLUSIONS

This thesis has offered a unified and integrated approach to the application of the state-of-the-art optimization techniques to circuit design, nominal as well as statistical, and modeling problems. Essential aspects related to both the formulation of the problems and an effective and efficient solution method have been addressed.

At the heart of our formulation of the design and modeling problems is a hierarchy of circuit models which unified the presentation of Chapters 4, 5 and 6. The use of idealized models has led to the optimization of a single set of nominal circuit parameters by minimizing a suitable  $\ell_p$  measure of the errors between the performance specifications and the circuit responses of interest. The explicit consideration of tolerances, model uncertainties as well as measurement inadequacy and inaccuracy has given rise to the more realistic multi-circuit approaches. In design centering, the multiple circuits are generated according to a suitable statistical assumption, whereas in modeling they are created from deliberate manipulations of the system. In both cases, our aim has been to expose and overcome the uncertainties that inevitably exist in an engineering problem.

In this thesis, gradient-based optimization techniques have not only served as powerful tools for solving abstract mathematical

problems, but also contributed as an integral part to the overall strategy. The extensive use of minimax optimization in design is justified by the fact that equal-ripple responses are both feasible and desirable for the filters, amplifiers and multiplexers that we have considered. Without exploiting the theoretical properties of the  $\ell_1$  norm, the multi-circuit modeling approach might not have been successful. The discontinuity in derivatives has in the past frustrated the application of the minimax and  $\ell_1$  objectives. The critical contributions made by Hald and Madsen have resulted in a class of fast and reliable algorithms for nonlinear  $\ell_p$  optimization as we have described in Chapter 2. These algorithms have demonstrated a superior performance in numerous circuit applications, including design, modeling, tuning and fault diagnosis.

The supply of derivatives, exact or approximate, is also an integral part of gradient-based circuit optimization. In Chapter 3, we have described efficient approaches to exact sensitivity calculation, applicable at both terminated network and unterminated subnetwork levels. When an exact sensitivity expression is not available, we have developed gradient approximations in Chapter 7 which effectively and efficiently integrate the optimization algorithm with the circuit simulation module. Since many commercial CAD programs currently available lack the capability of providing exact sensitivities, gradient approximations can and should be utilized to facilitate the practical implementation of advanced optimizers.

The theoretical results presented in this thesis have been

supported amply by circuit examples. These examples served both as illustrations of the feasibility and efficiency of the proposed algorithms and to maintain the engineering relevance of the work. Particularly, the applications to cavity filters, multiplexers and FET devices are of current significance and their development has been motivated by industrial demand. The close cooperation with industry has proved invaluable to the theoretical development. For instance, the model verification technique described in Chapter 6 was investigated only after some limitations of the experimental environment became evident as the actual data was processed. In return, the theoretical advances will undoubtedly further the industrial application of modern CAD techniques.

A number of problems related to the topics in this thesis are worth further research and development.

- (a) In Chapter 5, we have described a generalized  $\ell_p$  centering algorithm. Like most statistical design methods, its major computational effort is related to the Monte Carlo analyses of the circuit (i.e., simulation of the multiple circuits). Although we have proposed a simulation saving technique, it would be extremely useful to incorporate multidimensional approximations, which we have reviewed in Section 5.3.2, into the algorithm. By this approach, a suitable model, linear, maximally flat or quadratic, is constructed for each error function by performing exact simulations at a number of base points. During optimization, such models are used to provide function values and derivatives instead

of requiring new circuit simulations. This would improve the efficiency of the centering algorithm tremendously, especially for large networks.

(b) The multi-circuit modeling technique described in Chapter 6 can be combined with the functional tuning approach proposed by Bandler and Salama (1985a) to develop a strategy for computer-aided actual tuning (as opposed to simulated tuning). From measurements made on a manufactured device, the modeling technique is used to produce a consistent and reliable relationship between the physically adjustable elements and the parameters of an equivalent circuit. Such a relationship must be updated automatically and adaptively as the tuning proceeds. Using this result the algorithm would suggest the necessary adjustments on the physical device. Anticipating imprecisions in the implementation, the determination of tuning adjustments may have to be toleranced. In other words, such algorithms may have to integrate the relevant concepts of multi-circuit design and modeling.

(c) Another recent research direction in device modeling has been model evolution, i.e., automatic modification of the equivalent circuit topology. At the present time, such strategies are usually heuristic and of limited use. However, for specific applications, it may be possible and useful to combine a circuit topology modification technique with the modeling approach introduced in this thesis. For

instance, we can process simultaneously a simplified model using DC or low-frequency measurements and a more complicated model at the normal operating frequencies. Common variables can be identified between these models since they are supposed to represent the same device.

- (d) In Chapter 7, we have developed a weighted update for the gradient approximation algorithm. At the present time, the weights are defined prior to and kept constant during the optimization. The application of the weighted update may be broadened if we can develop an automatic and adaptive scheme for modifying the weights during the optimization.

## APPENDIX

### FORMULAS FOR POWELL'S SPECIAL ITERATIONS

The algorithm for computing the increment vector for a special iteration, as derived by Powell (1970a), is as follows.

An  $n$  by  $n$  ( $n$  being the dimension of  $\mathbf{x}$ ) orthogonal matrices  $\mathbf{D}_k$  is constructed at each iteration. Denote the rows of  $\mathbf{D}_k$  by  $\mathbf{d}_i^T$ ,  $i = 1, 2, \dots, n$ . At a special iteration, the increment vector is set to a multiple of the first row vector of  $\mathbf{D}_k$ , as

$$\mathbf{h}_k = \Lambda_k \mathbf{d}_1, \quad (\text{A.1})$$

where  $\Lambda_k$  is a parameter controlling the step size of  $\mathbf{h}_k$ . Usually it is set to the step size of the latest ordinary iteration.

At the starting point  $\mathbf{D}_1$  is set to an identity matrix. At the  $k$ th iteration  $\mathbf{D}_k$  is revised to produce  $\mathbf{D}_{k+1}$ . We use  $\mathbf{y}_i^T$  for the rows of  $\mathbf{D}_{k+1}$ . For a special iteration, we simply let

$$\begin{aligned} \mathbf{y}_i &= \mathbf{d}_{i+1}, \quad i = 1, 2, \dots, n-1, \\ \mathbf{y}_n &= \mathbf{d}_1. \end{aligned} \quad (\text{A.2})$$

For an ordinary iteration, the following steps take place.

Step 1 Compute  $\sigma_i = \mathbf{d}_i^T \mathbf{h}_k$ ,  $i = 1, 2, \dots, n$ .

Step 2 Find  $t$  which is the greatest integer such that  $\sigma_t \neq 0$ .

Step 3 Let  $\alpha_t = 0$  and  $\mathbf{z}_t = \mathbf{0}$ . For  $i = t-1, t-2, \dots, 1$ , compute

$$\begin{aligned} \mathbf{z}_i &= \mathbf{z}_{i+1} + \sigma_{i+1} \mathbf{d}_{i+1}, \\ \alpha_i &= \alpha_{i+1} + \sigma_{i+1}^2, \\ \mathbf{y}_i &= (\alpha_i \mathbf{d}_i - \sigma_i \mathbf{z}_i) / [\alpha_i (\alpha_i + \sigma_i^2)]^{\frac{1}{2}}. \end{aligned} \quad (\text{A.3})$$

Step 4 Let  $\mathbf{y}_i = \mathbf{d}_{i+1}$ ,  $i = t, t+1, \dots, n-1$ . Let  $\mathbf{y}_n = \mathbf{h}_k / (\mathbf{h}_k^T \mathbf{h}_k)^{\frac{1}{2}}$ .

## REFERENCES

- H.L. Abdel-Malek and J.W. Bandler (1980), "Yield optimization for arbitrary statistical distributions, part I: theory, part II: implementation", IEEE Trans. Circuits and Systems, vol. CAS-27, pp. 245-262.
- A.E. Atia (1974), "Computer-aided design of waveguide multiplexers", IEEE Trans. Microwave Theory Tech., vol. MTT-22, pp. 332-336.
- A.E. Atia and A.E. Williams (1971), "New types of waveguide bandpass filters for satellite transponders", COMSAT Technical Review, vol. 1, pp. 21-43.
- A.E. Atia and A.E. Williams (1972), "Narrow-bandpass waveguide filters", IEEE Trans. Microwave Theory Tech., vol. MTT-20, pp. 258-265.
- A.E. Atia and A.E. Williams (1974), "Nonminimum-phase optimum-amplitude bandpass waveguide filters", IEEE Trans. Microwave Theory Tech., vol. MTT-22, pp. 425-431.
- J.W. Bandler and H.L. Abdel-Malek (1978), "Optimal centering, tolerancing, and yield determination via updated approximations and cuts", IEEE Trans. Circuits and Systems, vol. CAS-25, pp. 853-871.
- J.W. Bandler and C. Charalambous (1972), "Theory of generalized least pth approximation", IEEE Trans. Circuit Theory, vol. CT-19, pp. 287-289.
- J.W. Bandler and C. Charalambous (1974), "Nonlinear programming using minimax techniques", J. Opt. Theory Applic., vol. 13, pp. 607-619.
- J.W. Bandler and S.H. Chen (1984a), "Interactive optimization of multi-coupled cavity microwave filters", Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, Report SOS-84-13.
- J.W. Bandler and S.H. Chen (1984b), "Case study of 10th order elliptic and quasi-elliptic self-equalized multi-cavity filters", Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, Report SOS-84-23.

J.W. Bandler, S.H. Chen and S. Daijavad (1984a), "Proof and extension of general sensitivity formulas for lossless two-ports", Electronics Letters, vol. 20, pp. 481-482.

J.W. Bandler, S.H. Chen and S. Daijavad (1984b), "Novel approach to multicoupled-cavity filter sensitivity and group delay computation", Electronics Letters, vol. 20, pp. 580-582.

J.W. Bandler, S.H. Chen and S. Daijavad (1985), "Simple derivation of a general sensitivity formula for lossless two-ports", Proc. IEEE, vol. 73, pp. 165-166.

J.W. Bandler, S.H. Chen and S. Daijavad (1986a), "Exact sensitivity analysis for optimization of multi-coupled cavity filters", Int. J. Circuit Theory and Applic., vol. 14, pp. 63-77.

J.W. Bandler, S.H. Chen and S. Daijavad (1986b), "Microwave device modeling using efficient  $\ell_1$  optimization: a novel approach", IEEE Trans. Microwave Theory Tech., vol. MTT-34, pp. 1282-1293.

J.W. Bandler, S.H. Chen, S. Daijavad and W. Kellermann (1984), "Optimal design of multi-cavity filters and contiguous-band multiplexers", Proc. 14th European Microwave Conf. (Liege, Belgium), pp. 863-868.

J.W. Bandler, S.H. Chen, S. Daijavad, W. Kellermann, M. Renault and Q.J. Zhang (1986), "Large scale minimax optimization of microwave multiplexers", Proc. 16th European Microwave Conf. (Dublin, Ireland), pp. 435-440.

J.W. Bandler, S.H. Chen, S. Daijavad and K. Madsen (1986), "Efficient gradient approximations for nonlinear optimization of circuits and systems", Proc. IEEE Int. Symp. Circuits and Systems (San Jose, CA), pp. 964-967.

J.W. Bandler, S.H. Chen, S. Daijavad and K. Madsen (1988), "Efficient optimization with integrated gradient approximations", IEEE Trans. Microwave Theory Tech., vol. MTT-36.

J.W. Bandler, S.H. Chen and M.L. Renault, "KMOS - a Fortran library for nonlinear optimization", Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, Report SOS-87-1.

J.W. Bandler, S. Daijavad and Q.J. Zhang (1986), "Exact simulation and sensitivity analysis of multiplexing networks", IEEE Trans. Microwave Theory Tech., vol. MTT-34, 1986, pp. 93-102.

- J.W. Bandler, M.A. El-Kady, W. Kellermann and W.M. Zuberek (1983), "An optimization approach to the best alignment of manufactured and operating systems", Proc. IEEE Int. Symp. Circuits and Systems (Newport Beach, CA), pp. 542-545.
- J.W. Bandler and W. Kellermann (1983), "Selected topics in optimal design centering, tolerancing and tuning", Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, Report SOS-83-28.
- J.W. Bandler, W. Kellermann and K. Madsen (1985), "A superlinearly convergent minimax algorithm for microwave circuit design", IEEE Trans. Microwave Theory Tech., vol. MTT-33, pp. 1519-1530.
- J.W. Bandler, W. Kellermann and K. Madsen (1987), "A nonlinear  $\ell_1$  optimization algorithm for design, modelling and diagnosis of networks", IEEE Trans. Circuits and Systems, vol. CAS-34, pp.174-181.
- J.W. Bandler, P.C. Liu and H. Tromp (1976a), "A nonlinear programming approach to optimal design centering, tolerancing and tuning", IEEE Trans. Circuits and Systems, vol. CAS-23, pp. 155-165.
- J.W. Bandler, P.C. Liu and H. Tromp (1976b), "Integrated approach to microwave design", IEEE Trans. Microwave Theory Tech., vol. MTT-24, pp. 584-591.
- J.W. Bandler and P.A. Macdonald (1969), "Cascaded noncommensurate transmission-line networks as optimization problems", IEEE Trans. Circuit Theory, vol. CT-16, pp. 391-394.
- J.W. Bandler and M.R.M. Rizk (1979), "Optimization of electrical circuits", Math. Programming Study, vol. 11, pp.1-64.
- J.W. Bandler and A.E. Salama (1985a), "Fault diagnosis of analog circuits", Proc. IEEE, vol. 73, pp. 1279-1325.
- J.W. Bandler and A.E. Salama (1985b), "Functional approach to microwave postproduction tuning", IEEE Trans. Microwave Theory Tech., vol. MTT-33, pp. 302-310.
- J.W. Bandler and Q.J. Zhang (1987), "An automatic decomposition technique for device modelling and large circuit design", IEEE Int. Microwave Symp. Digest (Las Vegas, NA), pp. 709-712.
- J.W. Bandler and W.M. Zuberek (1982), "MMLC - A Fortran package for linearly constrained minimax optimization", Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, Report SOS-82-5.
- R.H. Bartels and A.R. Conn (1981), "An approach to nonlinear  $\ell_1$  data fitting", Computer Science Department, University of Waterloo, Waterloo, Canada, Report CS-81-17.

- R.M. Biernacki and M.A. Styblinski (1986), "Statistical circuit design with a dynamic constraint approximation scheme", Proc. IEEE Int. Symp. Circuits and Systems (San Jose, CA), pp. 976-979.
- F.H. Branin (1973), Jr., "Network sensitivity and noise analysis simplified", IEEE Trans. Circuit Theory, vol. CT-20, pp. 285-288.
- R.K. Brayton, G.D. Hachtel and A.L. Sangiovanni-Vincentelli (1981), "A survey of optimization techniques for integrated-circuit design", Proc. IEEE, vol. 69, pp. 1334-1362.
- R.P. Brent (1972), "On the Davidenko-Branin method for solving simultaneous nonlinear equations", IBM J. Research and Development, vol. 16, pp. 434-436.
- C.G. Broyden (1965), "A class of methods for solving nonlinear simultaneous equations", Math. Comp., vol. 19, pp. 577-593.
- C.G. Broyden (1967), "Quasi-Newton methods and their application to function minimization", Math. Comp., vol. 21, pp. 368-381.
- C.G. Broyden (1969), "A new double-rank minimization algorithm", Notices American Math. Society, vol. 16, p. 670.
- E.M. Butler (1971), "Realistic design using large-change sensitivities and performance contours", IEEE Trans. Circuit Theory, vol. CT-18, pp. 58-66.
- R.J. Cameron (1982), "Dual-mode realisations for asymmetric filter characteristics", ESA Journal, vol. 6, pp. 339-356.
- C. Charalambous (1977), "Nonlinear least pth optimization and nonlinear programming", Math. Programming, vol. 12, pp. 195-225.
- M.H. Chen (1985), "Current state-of-the-art technology on contiguous-band multiplexer", Proc. IEEE Int. Symp. Circuits and Systems (Kyoto, Japan), pp. 1583-1586.
- M.H. Chen, F. Assal and C. Mahle (1976), "A contiguous band multiplexer", COMSAT Technical Review, vol. 6, pp. 285-306.
- F.H. Clarke (1975), "Generalized gradients and applications", Trans. American Math. Society, vol. 205, pp. 247-262.
- ComDev Ltd. (1985), 155 Sheldon Drive, Cambridge, Ontario, Canada N1R 7H6, private communications.
- ComDev Ltd. (1986), 155 Sheldon Drive, Cambridge, Ontario, Canada N1R 7H6, private communications.

W.R. Curtice and R.L. Camisa (1984), "Self-consistent GaAs FET models for amplifier design and device diagnostics", IEEE Trans. Microwave Theory Tech., vol. MTT-32, pp. 1573-1578.

S. Daijavad (1986), "Design and modelling of microwave circuits using optimization methods", Ph.D. Thesis, McMaster University, Hamilton, Canada.

G.B. Dantzig, "Maximization of a linear function of variables subject to linear inequalities", in Activity Analysis of Production and Allocation, Cowles Commission Monograph 13, T.C. Koopmans, ed., John Wiley, New York.

W.C. Davidon (1959), "Variable metric method for minimization", Rep. ANL-5990 Rev, Argonne National Laboratories, Argonne, IL.

J.E. Dennis, Jr. and J.J. Moré (1977), "Quasi-Newton methods, motivation and theory", SIAM Review, vol. 19, pp. 46-89.

S.W. Director and G.D. Hachtel (1977), "The simplicial approximation approach to design centering", IEEE Trans. Circuits and Systems, vol. CAS-24, pp. 363-372.

S.W. Director and R.A. Rohrer (1969a), "Generalized adjoint network and network sensitivities", IEEE Trans. Circuit Theory, vol. CT-16, pp. 318-323.

S.W. Director and R.A. Rohrer (1969b), "Automated network design: The frequency domain case", IEEE Trans. Circuit Theory, vol. CT-16, pp. 330-337.

L.C.W. Dixon (1972), "Quasi-Newton algorithms generate identical points", Math. Programming, pp. 383-387.

R.A. El-Attar, M. Vidyasagar and S.R.K. Dutta, "An algorithm for  $\ell_1$ -norm minimization with application to nonlinear  $\ell_1$ -approximation", SIAM J. Numerical Analysis, vol. 16, pp. 70-86.

N.J. Elias (1975), "New statistical methods for assigning device tolerances", Proc. IEEE Int. Symp. Circuits and Systems (Newton, MA), pp. 329-332.

R. Fletcher (1970), "A new approach to variable metric algorithms", Comput. J., vol. 13, pp. 317-322.

R. Fletcher and M.J.D. Powell (1963), "A rapidly convergent descent method for minimization", Comput. J., vol. 6, pp. 163-168.

D. Goldfarb (1970), "A family of variable-metric methods derived by variational means", Math. Comp., vol. 24, pp. 23-26.

- G.D. Hachtel and A.L. Sangiovanni-Vincentelli (1981), "A survey of third-generation simulation techniques", Proc. IEEE, vol. 69, pp. 1264-1280.
- J. Hald and K. Madsen (1981), "Combined LP and quasi-Newton methods for minimax optimization", Math. Programming, vol. 20, pp. 49-62.
- J. Hald and K. Madsen (1985), "Combined LP and quasi-Newton methods for nonlinear  $l_1$  optimization", SIAM J. Numerical Analysis, vol. 22, pp. 68-80.
- B.J. Karafin (1971), "The optimum assignment of component tolerances for electrical networks", The Bell System Technical Journal, vol. 50, pp. 1225-1242.
- C.M. Kudsia (1982), "Manifestations and limits of dual-mode filter configurations for communications satellite multiplexers", AIAA 9th Comm. Satellite Systems Conf. (San Diego, CA), pp. 294-303.
- H.W. Kuhn and A.W. Tucker (1951), "Non-linear programming", Proc. 2nd Symp. on Math. Statistics and Probability (Berkeley, CA), University of California Press, pp. 481-493.
- K.R. Laker, M.S. Ghausi and J.J. Kelly (1975), "Minimum sensitivity active (leapfrog) and passive ladder bandpass filters", IEEE Trans. Circuits and Systems, vol. CAS-22, pp. 670-677.
- S. Lang (1977), Complex Analysis, Addison-Wesley, Reading, MA, p. 32.
- L.S. Lasdon, D.F. Suchman and A.D. Waren (1966), "Nonlinear programming applied to linear array design", J. Acoust. Soc. Am., vol. 40, pp. 1197-1200.
- K. Levenberg (1944), "A method for the solution of certain problems in least squares", Quart. Appl. Math., vol. 2, pp. 164-168.
- K. Madsen (1975), "Minimax solution of nonlinear equations without calculating derivatives", Math. Programming Study, vol. 3, pp. 110-126.
- K. Madsen (1985), "Minimization of non-linear approximation functions", Dr. techn. thesis, Institute of Numerical Analysis, Tech. Univ. of Denmark, DK2800 Lyngby, Denmark.
- K. Madsen (1986), Institute of Numerical Analysis, Tech. Univ. of Denmark, DK2800 Lyngby, Denmark, Unpublished algorithm for least squares, private communications.
- K. Madsen and H. Schjaer-Jacobsen (1978), "New algorithms for worst case tolerance optimization", Proc. IEEE Int. Symp. Circuits and Systems (New York), pp. 681-685.

- D.W. Marquardt (1963), "An algorithm for least-squares estimation of nonlinear parameters", SIAM J. Appl. Math., vol. 11, pp. 431-441.
- J.J. Moré (1982), "Recent developments in algorithms and software for trust region methods", in Mathematical Programming, the state of the art (Bonn 1982), Springer Verlag, pp. 258-287.
- H.J. Orchard, G.C. Temes and T. Cataltepe (1983), "General sensitivity formulas for terminated lossless two-ports", Electronics Letters, vol. 19, pp. 576-578.
- H.J. Orchard, G.C. Temes and T. Cataltepe (1985), "Sensitivity formulas for terminated lossless two-ports", IEEE Trans. Circuits and Systems, vol. CAS-32, pp. 459-466.
- M.R. Osborne and G.A. Watson (1969), "An algorithm for minimax optimization in the nonlinear case", Comput. J., vol. 12, pp.63-68.
- M.R. Osborne and G.A. Watson (1971), "On an algorithm for discrete nonlinear  $\ell_1$  approximation", Comput. J., vol. 14, pp. 184-188.
- J.F. Pinel and K.A. Roberts (1972), "Tolerance assignment in linear networks using nonlinear programming", IEEE Trans. Circuit Theory, vol. CT-19, pp. 475-479.
- E. Polak and A.L. Sangiovanni-Vincentelli (1979), "Theoretical and computational aspects of the optimal design centering, tolerancing, and tuning problem", IEEE Trans. Circuits and Systems, vol. CAS-26, pp. 795-813.
- M.J.D. Powell (1970a), "A Fortran subroutine for unconstrained minimization, requiring first derivatives of the objective functions", Atomic Energy Research Establishment, Harwell, Berkshire, England, Report AERE-R. 6469, pp. 20-27.
- M.J.D. Powell (1970b), "A new algorithm for unconstrained optimization", in Nonlinear Programming, J.B. Rosen, O.L. Mangasarian and K. Ritter, eds., Academic Press, New York.
- M.J.D. Powell (1970c), "A hybrid method for nonlinear equations", in Numerical Methods for Nonlinear Algebraic Equations, P. Rabinowitz, ed., Gordon and Breach, London.
- M.J.D. Powell (1978), "The convergence of variable metric methods for non-linearly constrained optimization calculations", in Nonlinear Programming 3, O.L. Mangasarian, R.R. Meyer and S.M. Robinson, eds., Academic Press, New York, pp. 27-63.
- R.A. Pucel (1986), Raytheon Company, Research Division, Lexington, MA 02173, FET data, private communications.

- J. Schoeffler (1964), "The synthesis of minimum sensitivity networks", IEEE Trans. Circuit Theory, vol. CT-11, pp. 271-276.
- D.F. Shanno (1970), "Conditioning of quasi-Newton methods for function minimization", Math. Comp., vol. 24, pp. 647-656.
- K. Singhal and J.F. Pinel (1981), "Statistical design centering and tolerancing using parametric sampling", IEEE Trans. Circuits and Systems, vol. CAS-28, pp. 692-701.
- R.S. Soin and R. Spence (1980), "Statistical exploration approach to design centering", IEE. Proc., vol.127, Pt. G., pp.260-269.
- SUPER-COMPACT User's Manual (1986), Communications Consulting Corp., Patterson, NJ 07504, May 1986.
- G.C. Temes and D.A. Calahan (1967), "Computer-aided network optimization the state-of-the-art", Proc. IEEE, vol. 55, pp. 1832-1863.
- G.C. Temes and D.Y.F. Zai (1969), "Least pth approximation", IEEE Trans. Circuit Theory, vol. CT-16, pp. 235-237.
- TOUCHSTONE User's Manual (1985), EEsof Inc., Westlake Village, CA 91362.
- TOUCHSTONE Version 1.5 (1987), EEsof Inc., Westlake Village, CA 91362.
- H. Tromp (1977), "The generalized tolerance problem and worst case search", Conf. Computer-aided Design of Electronic and Microwave Circuits and Systems (Hull, England), pp. 72-77.
- H. Tromp (1978), "Generalized worst case design, with applications to microwave networks", Doctoral Thesis (in Dutch), Faculty of Engineering, University of Ghent, Ghent, Belgium.
- A.D. Waren, L.S. Lasdon and D.F. Suchman (1967), "Optimization in engineering design", Proc. IEEE, vol. 55, pp. 1885-1897.
- W.M. Zuberek (1984), "Numerical approximation of gradients for circuit optimization", Proc. 27th Midwest Symp. Circuits and Systems (Morgantown, WV), pp. 200-203.

## AUTHOR INDEX

H.L. Abdel-Malek	97, 99
F. Assal	58, 77, 80
A.E. Atia	58, 59, 60, 67, 77
J.W. Bandler	7, 11, 19, 21, 22, 27, 29, 33, 40, 46, 49, 50, 51, 52, 53, 58, 65, 67, 77, 80, 81, 82, 87, 92, 97, 98, 99, 116, 117, 120, 125, 129, 132, 135, 151, 161, 165, 175, 193, 194, 203
R.H. Bartels	10, 125
R.M. Biernacki	97, 99
F.H. Branin, Jr.	34
R.K. Brayton	108
R.P. Brent	165
C.G. Broyden	5, 17, 20, 151, 153, 194
E.M. Butler	85, 92
D.A. Calahan	1, 48
R.J. Cameron	58, 70
R.L. Camisa	134
T. Cataltepe	33, 40
C. Charalambous	7, 11, 18, 98
M.H. Chen	58, 77, 80
S.H. Chen	29, 33, 40, 49, 58, 65, 67, 77, 80, 116, 120, 125, 129, 132, 151

F.H. Clarke	19
A.R. Conn	10, 125
W.R. Curtice	134
S. Daijavad	33, 40, 46, 49, 58, 67, 77, 80, 116, 120, 125, 129, 132, 146, 151, 174
G.B. Dantzig	12
W.C. Davidon	17, 20
J.E. Dennis	20
S.W. Director	32, 34, 92, 98
L.C.W. Dixon	20
S.R.K. Dutta	193
R.A. El-Attar	193
M.A. El-Kady	50
N.J. Elias	85, 92
R. Fletcher	17, 20, 65
M.S. Ghausi	85
D. Goldfarb	20
G.D. Hachtel	46, 97, 98, 108
J. Hald	3, 7, 8, 16, 19, 21, 23, 31, 160, 201
B.J. Karafin	85, 92, 187
W. Kellermann	19, 21, 22, 27, 30, 50, 77, 80, 81, 92, 117, 135, 165, 174, 193, 194
J.J. Kelly	85
C.M. Kudsia	58, 64
H.W. Kuhn	8, 18
K.R. Laker	85

S. Lang	45
L.S. Lasdon	48
K. Levenberg	17
P.C. Liu	51, 52, 87, 92, 97
P.A. Macdonald	161
K. Madsen	3, 7, 8, 15, 16, 19, 21, 22, 23, 27, 29, 30, 31, 80, 81, 117, 135, 151, 154, 160, 161, 165, 174, 179, 187, 188, 193, 194, 201
C. Mahle	58, 77, 80
D.W. Marquardt	17
J.J. Moré	16, 20
H.J. Orchard	33, 40
M.R. Osborne	16
J.F. Pinel	85, 92, 95, 97, 100, 109, 110, 112
E. Polak	97, 98
M.J.D. Powell	5, 17, 20, 21, 27, 65, 151, 154, 157, 205
R.A. Pucel	134
M.L. Renault	29, 80
M.R.M. Rizk	53
K.A. Roberts	85, 92
R.A. Rohrer	32, 34
A.E. Salama	126, 203
A.L. Sangiovanni-Vincentelli	46, 97, 98, 108
H. Schjaer-Jacobsen	161
J. Schoeffler	85

D.F. Shanno	20
K. Singhal	92, 95, 97, 100, 109, 110, 112
R.S. Soin	97, 100
R. Spence	97, 100
M.A. Styblinski	97, 99
D.F. Suchman	48
G.C. Temes	1, 7, 9, 33, 40, 48
H. Tromp	50, 51, 52, 87, 92, 97
A.W. Tucker	8, 18
M. Vidyasagar	193
A.D. Waren	48
G.A. Watson	16
A.E. Williams	58, 59, 60, 67
D.Y.F. Zai	7, 9
Q.J. Zhang	46, 80, 82, 174
W.M. Zuberek	29, 50, 151, 154, 179

## SUBJECT INDEX

- Acceptable region, 98
- Active set, 19, 26, 28
- Adjoint network, 32, 34, 38
- Admittance matrix, 33, 42, 61
- Amplifier, 169
- Approximations,
  - least pth, 7
  - maximally flat, 99
  - multidimensional polynomial, 99
  - simplicial, 98
- Broyden update, 153
- Combined methods, 21, 28
- Common variables, 120
- Cost functions, 97, 107
- Cubic interpolation, 65
- Decomposition, 46, 80
- Design,
  - centering, 91, 102
  - nominal, 48, 57
  - multi-circuit, 88
  - statistical, 85, 92
- Dissipation, 59, 74
- Error functions, 54, 118
- Fault diagnosis and location, 194
- FET,
  - equivalent circuit, 132
  - modeling, 132, 196
- Gauss-Newton methods, 14, 21, 23

- Generalized  $\ell_p$ ,
  - centering, 102
  - functions, 11, 103
- Gradient approximations,
  - by perturbations, 152
  - initial, 157
  - algorithms, 150, 157, 160, 187
- Gravity method, 100
- Group delay, 39, 41, 67
- Hessians, 18, 19, 151, 155
- Identifiability,
  - measure, 126
  - necessary conditions, 128, 129
- Jacobians, 15, 18, 27, 126, 153
- Kuhn-Tucker conditions, 18, 25
- Large-scale networks, 46, 77
- $\ell_1$ ,
  - algorithms, 29, 187
  - norm, 10
  - optimality, 18
  - properties, 10, 125
- $\ell_p$ ,
  - norms, 9
  - objective functions, 58, 118, 121
- Least squares, 9, 17, 29
- Linear  $\ell_p$  problems, 12
- Linear programming, 12, 23
- Linearization, 14, 24
- Linear constraints, 14, 22
- Minimax,
  - algorithms, 29, 160
  - norm, 10
  - design, 59, 65, 103, 169, 171

- Modeling,
  - multi-circuit, 116, 118
  - single-circuit, 117
  - tuning-oriented, 142, 203
  - uniqueness, 118, 126
  - verification, 135
  
- Multi-coupled cavity filters,
  - centering, 111
  - equivalent circuit, 59
  - modeling, 140, 145
  - nominal design, 66, 70, 164
  - sensitivity, 61
  - structures, 59, 64
  
- Multiple circuits, 88, 92, 120
  
- Multi-ports, 34
  
- Multiplexers, 77, 171
  
- Nominal circuit, 57, 88
  
- Nonlinear programming, 11, 25, 97
  
- One-sided  $\ell_1$ ,
  - algorithm, 22, 27, 105
  - function, 22, 105
  - optimality, 26
  
- One-sided  $\ell_p$  functions, 10
  
- Optimality equations, 18, 26
  
- Parametric sampling, 100
  
- Parasitic effects, 87, 119, 135, 140
  
- Physical parameters, 49, 86, 142
  
- Quadratic programming, 14, 17
  
- Quasi-Newton methods, 17, 25
  
- Sensitivity,
  - frequency responses, 40
  - multi-coupled cavity filters, 61
  - multi-ports, 35
  - nodal description, 33
  - second-order, 39, 63

- two-ports, 38
- Simulation model,
  - hierarchy, 49, 86, 118
  - parameters, 51
  - responses, 51
- Simulation saving techniques, 111, 202
- Special iterations of Powell, 157, 205
- Specifications, 53
- Tolerances, 86, 89, 97
- Tolerancing, 91, 107
- Transformers,
  - input and output, 67, 140, 180
  - stripline, 51, 87
  - transmission line, 161, 165, 188
- Trust region, 15, 21, 23
- Tuning, 91, 97, 102, 107, 203
- Two-ports, 38, 40, 61
- Uncertainties, 87, 118
- Weighted update, 155, 182, 194, 197, 204
- Worst-case design, 97, 103, 169
- Yield,
  - estimation, 89, 101
  - enhancement, 91, 105