ADVANCES IN YIELD-DRIVEN DESIGN OF MICROWAVE CIRCUITS

By

JIAN SONG, B. Eng., M. Eng.
(Chongqing University)

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

April 1991

ADVANCES IN YIELD-DRIVEN DESIGN

DOCTOR OF PHILOSOPHY(1991) McMASTER UNIVERSITY
(Electrical Engineering) Hamilton, Ontario

TITLE: Advances in Yield-Driven Design of Microwave Circuits

AUTHOR: Jian Song
B.Eng.(E.E.), M.Eng.(E.E.), (Chongqing University)

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: xiii, 163

ABSTRACT

This thesis addresses itself to computer-aided yield-driven design of microwave circuits using implementable, efficient approximation and optimization techniques.

Basic concepts of yield-driven design are identified. A number of approaches to statistical design are reviewed. Their features and limitations are discussed. The recent generalized ℓ_p centering approach and one-sided ℓ_1 optimization algorithm are addressed.

A highly efficient quadratic approximation, specially applicable to statistical design, is presented. A set of very simple and easy-to-implement formulas is derived. This approximation technique is also applied to gradient functions of circuit responses to provide higher accuracy.

A combined approach to attack large scale problem is presented, which explores the most powerful capabilities of hardware and software available to us, namely, the supercomputer, efficient quadratic modeling, fast and dedicated simulation, and state-of-the-art optimization. Yield-driven design techniques are extended to deal with tunable circuits by considering tuning tolerances. A 5-channel waveguide multiplexer is considered as an example both for the combined approach and for the treatment of tunable circuits.

Yield-driven design of nonlinear microwave circuits with statistically characterized devices is considered. Relevant concepts are introduced. The efficient Integrated Gradient Approximation Technique (*IGAT*) is presented in the statistical

design environment, which avoids the prohibitive computational burden resulting from the traditional perturbation scheme.

A novel approach, called Feasible Adjoint Sensitivity Technique (*FAST*), is derived to calculate sensitivities of nonlinear circuits that are simulated in the harmonic balance environment. By taking advantage of the computational efficiency of adjoint analysis and the implementational simplicity of the perturbation technique, *FAST* is responsible for great savings of computational effort required for yield-driven design of nonlinear circuits.

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. J.F. MacGregor and Dr. N.K. Sinha, members of his Supervisory Committee, for their valuable suggestions and continuing interest.

The author appreciates the opportunity given to him by Dr. Bandler to be involved in industrially relevant projects and academic research activities. Cooperation between Dr. Bandler and some leading industrial establishments has always been a significant motivation for the work on yield-driven design.

The author offers his deep gratitude to Dr. R.M. Biernacki of McMaster University and Optimization Systems Associates Inc. for his expert advice, through frequent and very beneficial discussions.

It is the author's pleasure to have been associated with Dr. S.H. Chen and Dr. Q.J. Zhang. The continuous and in-depth interaction with them has led to the development and improvement of new ideas. Some of their excellent work has been very helpful in the author's research. The author has also benefitted from inspiring discussions with his colleagues S. Ye and Q. Cai.

The assistance of M.L. Renault and G.R. Simpson of the Simulation Optimization Systems Research Laboratory (SOSRL) in preparing programs and providing computer and word processing facilities is acknowledged.

The financial assistance provided by the Natural Sciences and Engineering Research Council of Canada through Grants OGP0007239, STR0040923,

EQP0043573, and the Department of Electrical and Computer Engineering through a Teaching Assistantship is gratefully acknowledged.

Finally, thanks are due to my family for their continuous encouragement, especially my wife, Shuxi Li, for showing her understanding, support and tolerance while she works towards her own Ph.D. degree.

TABLE OF CONTENTS

TABLE OF CONTENTS(continued)

TABLE OF CONTENTS(continued)

TABLE OF CONTENTS(continued)

LIST OF FIGURES

LIST OF TABLES

1

INTRODUCTION

For more than two decades, we have been witnesses to the rapid progress of circuit computer-aided design (CAD), from a set of abstract mathematic formulas understood by only a half dozen university professors to an indispensable and everyday tool used by virtually all circuit designers. Progress on such a scale is the result of two major driving forces, namely, active research and development of numerical circuit analysis and optimization, and the drastic evolution of computer hardware and software. CAD techniques continue to thrive, penetrating all circuit design areas with revolutionary ideas.

The monolithic microwave integrated circuit (MMIC) is part of a developing technology. The major difference between MMIC and its previous generation, the hybrid microwave integrated circuit (MIC), is the following: MMICs allow various active and passive circuit elements, such as transmission lines, resistors, capacitors, inductors, field-effect transistors (FETs) and diodes of many types to be integrated on a single chip, performing certain functions, such as amplification, mixing, filtering and phase-shifting. The MIC technology, however, is to mount active circuit elements and other components, for example, chip capacitors, on a dielectric substrate with previously fabricated transmission lines and other elements. In MIC engineering, tuning is possible by changing the physical dimensions of elements or even by substituting one FET for another. But for MMICs, there is no opportunity for device replacement and very limited scope for circuit tuning.

The traditional "performance-driven" design is considered adequate enough for MICs. This is because of the ability to replace devices and to tune circuits to counteract tolerance effects. The final production yield after tuning can be dramatically increased from the yield before tuning. However, the tuning procedure based on an individual circuit is very expensive and time consuming. It means that at the expense of higher costs, the final yield can be improved by tuning.

Because there is essentially no tuning allowed in MMIC production, imposing performance-driven design can lead to very low yield even for quite modest circuits and, consequently, to prohibitive costs. A different design approach, called "yield-driven" design or statistical design, is necessary for MMIC. Yield-driven design takes uncertainties, such as process tolerances, environmental fluctuations, model inaccuracy, etc., into account to maximize manufacturing yield.

Compared with digital circuit simulation, microwave circuit simulation is far more involved and complex because of the analog nature of such circuits. Microwave or analog circuits can have many different types of elements and most of them usually have continuous value distributions. Circuit responses are also continuous functions which demand considerable computational effort to solve. Simulation of nonlinear circuit responses is even more computationally intensive since it requires iteration. In yield-driven design, multiple sets of circuits, sometimes, a quite large number of them, need be simulated.

Optimization techniques are often utilized to automate yield-driven design. Because they are iterative in nature, most optimization algorithms require many objective function evaluations, each of which is the result of many circuit simulations. Gradient-based optimization techniques demonstrate, in general, far superior

performance to direct (non-gradient) methods. On the other hand, gradient calculation involves circuit sensitivity analysis.

As a result of many optimization iterations, multiple circuit simulations and sensitivity analyses, and possible iterative circuit simulation procedures, the collective computational effort creates special difficulties in terms of prohibitive computational costs and lengthy design cycles. To materialize the yield-driven design methodology, there is still a compelling need to further research and develop computationally feasible techniques to such an extent that yield-driven design is no longer merely an option, but an indispensable and robust tool for circuit designers dealing with practical circuits.

This thesis is intended to summarize new research results of computer-aided yield-driven design for microwave circuits. We propose new approaches to push yield optimization techniques forward to meet the challenges. The thesis consists of 7 chapters.

In Chapter 2, we identify concepts and notation involved in yield-driven design. Then, we review some of the most representative approaches developed in the past and the most recent approaches. Our emphasis is laid on the approach by Bandler and Chen (1989) since it will be used and further developed in succeeding chapters. The general formulation of this approach, which converts the problem of yield-driven design to an ℓ_1 optimization problem, is given. We consider the relevant concepts and definitions used in the one-sided ℓ_1 optimization. A two-stage algorithm to solve the one-sided ℓ_1 optimization problem is described.

In Chapter 3, we present a highly efficient quadratic approximation technique. The new approach takes advantage of the maximally flat interpolation and of a fixed

pattern of base points, thus substantially reducing computational effort and required storage. A set of extremely simple formulas to calculate model coefficients is derived. Moreover, this approach is extended to approximate gradients of circuit response functions. The elegance of this approach is its conciseness and applicability. High efficiency and feasibility for yield-driven design are demonstrated by the results of yield-driven circuit design.

Chapter 4 deals with statistical design of tunable circuits with tuning tolerances. The issue of statistical design of large scale circuits is also addressed. We propose a combined approach integrating the use of supercomputer, efficient quadratic modeling, and dedicated simulator. Modern supercomputers have found valuable applications in microwave circuit CAD with attractive performance-to-cost ratios. Our software, which carries out statistical design of microwave multiplexers, has been developed for the supercomputer. The computational results of a 5-channel multiplexer design performance on the Cray X-MP are reported in the chapter.

Chapter 5 offers the first comprehensive demonstration of yield-driven design of microwave circuits with statistically characterized devices. Efficient harmonic balance simulation is explored. A powerful gradient approximation technique, called *IGAT*, is introduced to avoid extremely expensive computational effort required by the traditional perturbation method. Large-signal FET parameter statistics are fully facilitated. Extensive numerical experiment directed at yield-driven design of a FET frequency doubler verifies our approach.

Chapter 6 presents a high-speed gradient calculation technique for microwave nonlinear circuits operating within the harmonic balance simulation environment. The technique, called *FAST*, is implementationally feasible for truly general-purpose

nonlinear circuit CAD, combining the efficiency and accuracy of the adjoint sensitivity analysis with the simplicity of the perturbation method. Through numerical experiments, the substantial advantages of *FAST* over other gradient calculation approaches have been observed.

We conclude in Chapter 7 with some suggestions for future research.

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

(1) An efficient quadratic approximation technique for statistical design.

(2) Extension of yield-driven design to tunable circuits with tuning tolerances.

(3) A combined approach using supercomputers, response approximation, and dedicated simulator, to statistical design of large scale circuits.

(4) A first integrated treatment of yield-driven design for nonlinear microwave circuits in the harmonic balance environment.

(5) A gradient approximation approach suitable for statistical design.

(6) A new, efficient, and easy-to-implement sensitivity analysis technique of nonlinear microwave circuits and its application in statistical design.

2

REVIEW OF STATISTICAL DESIGN APPROACHES AND OPTIMIZATION TECHNIQUES

2.1 INTRODUCTION

The use of optimization techniques can be traced back to the very first development of circuit CAD. Since then, due to design of more and more complex circuits and the enormous computational power available, the increasing demands have made optimization techniques widely adopted as primary tools in circuit CAD programs.

The traditional circuit design is the so-called nominal design that focuses on the improvement of individual circuit responses of interest. In reality, however, it is impossible to manufacture a circuit with exact designed parameters for the following reasons: the existence of uncertainties and tolerances inherent in the manufacturing process, the inaccuracy of mathematical models to approximate the real physical behaviour of circuit elements, etc. Recognition such fluctuations in the manufacturing environment and design process leads to the approach of yield-driven circuit design. The nature of yield-driven design suggests that it can be converted to an optimization problem, where the objective is no longer to improve individual circuit performance, but to increase the estimated manufacturing yield. Yield-driven design has been given much attention for more than two decades. Many approaches have been derived. There is a wide range of literature. A special issue of the IEEE Transactions on Computer-Aided Design on Statistical Design of VLSI Circuits Edited

by Strojwas and Sangiovanni-Vincentelli (1986) was published. There have been many CAD books and dedicated chapters on this subject (Strojwas 1987, Soin and Spence 1988, Ladbrooke 1989, and Vendelin, Pavio and Rohde 1990).

First, in this chapter, we introduce notation for circuit parameters, design variables, and their statistics. Circuit simulation and response calculation are addressed. Relations between response functions, design specifications, as well as error functions are discussed. The concepts of circuit outcomes, nominal values, tolerances, and manufacturing yield are identified. Then, the formal description of yield-driven design is presented. Various problems arising from statistical design are also described.

A review of a number of approaches to yield-driven design is given. We describe in some detail several methods which represent many years research on this subject.

Finally, we present the generalized ℓ_p centering approach (Bandler and Chen 1989) and elaborate on a special case: the ℓ_1 approach. We concentrate in detail on the one-sided ℓ_1 optimization problem and its implementational aspects because it will be used in our approach to yield-driven design. The basic concepts, relevant definitions, and mathematical formulation of the ℓ_1 optimization problem are presented. A two-stage algorithm combining the trust Gauss-Newton method and the quasi-Newton method is outlined.

2.2 NOTATION AND FORMULATIONS

In order to describe the formulation for yield-driven design, we will give relevant definitions and introduce useful notation.

2.2.1 Circuit Parameters, Design Variables and Their Statistics

The circuit considered will have a fixed topology and known component types, that is, the structure of the circuit is not an object of design and the components involved have the proper mathematical models to approximate their behaviour. A given set of parameter values will determine the performance of the circuit. The parameter values may be traditional discrete elements, such as resistors and capacitors, device parameters such as coefficients of characteristic equations of a FET model, geometrical dimensions such as the width of a microstrip line and the gate length of a FET, as well as manufacturing parameters such as the permeability and conductivity of the material. Other controlling operational and environmental factors, such as bias voltages and the excitation level, can also be included. We use

$$\phi \triangleq \begin{bmatrix} \phi_1 \\ \phi_2 \\ \cdot \\ \cdot \\ \cdot \\ \phi_N \end{bmatrix} \tag{2.1}$$

to denote this set of values where subscript N is the total number of parameters.

Relationships between these parameters can be very complicated. (1) A hierarchical structure among the parameters may exist. For example, the actual process parameters are defined, however, the circuit is simulated using equivalent

circuit models. A model is needed to convert process level parameters to the equivalent circuit parameters. (2) Some of parameters may be fixed or discrete for certain reasons. (3) Some of the parameters may be correlated because they are controlled by a group of lower level parameters. For instance, a number of equivalent circuit parameters are controlled by a group of common process parameters.

In a practical design process, only some of the ϕ are considered as *designable*. However, we shall assume that ϕ are all design variables unless otherwise stated. It is also physically meaningful to assume that design variables are independent.

In the classical design problem we are interested in finding one single point (circuit) in the design variable space which satisfies the design specifications. This kind of solution is impractical from the manufacturing point of view since there is a number of factors which influences the performance of a manufactured design. Among these factors are: (a) manufacturing tolerances, (b) model uncertainties, (c) parasitic effects, and (d) environment fluctuations.

Due to the foregoing uncertainties, the parameters of actually manufactured circuits are spread over a region. In the following, we use ϕ to denote a manufactured circuit. Suppose the probability density function (pdf) of ϕ is $p(\phi)$. In addition to the generic function form, two sets of parameters are needed to describe the probability density function. The first set ϕ^0 contains the nominal values, also called the nominal design or nominal circuit in circuit design. The second set of parameters can be tolerance extremes for uniform distribution, standard deviations associated with normal distributions, and correlation coefficients for joint distributions. We use

$$\epsilon \triangleq \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ . \\ . \\ . \end{bmatrix}$$

to denote this set of additional parameters. ϵ may have a higher dimension than ϕ since more parameters are needed to describe asymmetrical and correlated distributions. Now, the augmented form of the probability density function of ϕ becomes

$$p(\phi, \phi^0, \epsilon). \tag{2.2}$$

The tolerance region is determined by the probability density function and circuit outcomes are always fall into it

$$\phi \in R_\epsilon(\phi^0, \epsilon), \tag{2.3}$$

where $R_\epsilon(\phi^0, \epsilon)$ is the tolerance region. In the following, we may use $p(\phi)$, $p(\phi, \phi^0)$ or $p(\phi, \phi^0, \epsilon)$ to denote the probability density function according to different circumstances.

2.2.2 Circuit Simulation

Circuit response calculation usually involves a two-stage process. First, the circuit is simulated by solving a set of circuit equations

$$p(z, \phi) = 0, \tag{2.4}$$

where z, the solution of the equations, is a vector usually consisting of node voltages, branch voltages, branch currents, etc. In general, solving (2.4) involves systematic approaches to circuit analysis (Chua and Lin 1975, Vlach and Singhal 1983). This set

of circuit equations can be linear, nonlinear, differential-integral, or mixed. The circuit equations can also be based on nodal, loop, tabular, mixed analysis, etc.

Then, responses of design interest are calculated from the simulation results of (2.4). The responses are denoted by

$$F(\phi) \triangleq \begin{bmatrix} F_1(\phi) \\ F_2(\phi) \\ \cdot \\ \cdot \\ \cdot \\ F_M(\phi) \end{bmatrix}, \tag{2.5}$$

where M is the number of responses considered. In addition to ϕ, $F(\phi)$ is a function of z, which is intentionally omitted from $F(\phi)$ to avoid complexity.

Consider a voltage amplifier circuit. The response of interest is the voltage gain G_v. First, a set of nodal equations can be set up

$$Y(\phi)V = I, \tag{2.6}$$

where $Y(\phi)$ is the nodal admittance matrix which is assumed to exist, V the nodal voltage vector and I the nodal current excitation vector. We need to solve

$$p(V, \phi) = Y(\phi)V - I = 0. \tag{2.7}$$

Then, G_v is calculated from the voltage at the output node and the input voltage.

2.2.3 Design Specifications and Error Functions

In circuit design, specifications are used to describe the desired performance. Specifications are usually given in a set of discrete values. These values are functions of independent parameters, such as frequency, time, temperature, input power level, etc. (Bandler and Rizk 1979, Bandler, Biernacki, Chen, Song, Ye and Zhang 1990 and

1991a). We use ψ to denote the independent parameters. Specifications, denoted by $S(\psi)$, are arranged as the vector

$$S(\psi) \triangleq \begin{bmatrix} S_1(\psi) \\ S_2(\psi) \\ \cdot \\ \cdot \\ \cdot \\ S_L(\psi) \end{bmatrix}, \tag{2.8}$$

where L is the total number of specifications imposed. For simplicity, if we omit ψ in (2.8), then

$$S \triangleq \begin{bmatrix} S_1 \\ S_2 \\ \cdot \\ \cdot \\ \cdot \\ S_L \end{bmatrix}. \tag{2.9}$$

Specifications can be of the following forms: (1) an upper specification which requires the response be below it, (2) a lower specification which requires the response be above it, and (3) upper and lower specifications which require the same response be between the two. To distinguish upper and lower specifications, we introduce subscripts u and l, that is, S_{uj} and S_{lj}. For example, an amplifier is to be designed to meet a set of predetermined frequency specifications and, at the same time, to exhibit stable responses in a particular temperature region (Bandler and Rizk, 1979). A typical graphical presentation of such a case is shown in Fig 2.1. Two independent parameters, temperature and frequency, are involved to define the specifications.

For the case of upper or lower specifications, we define the error functions as

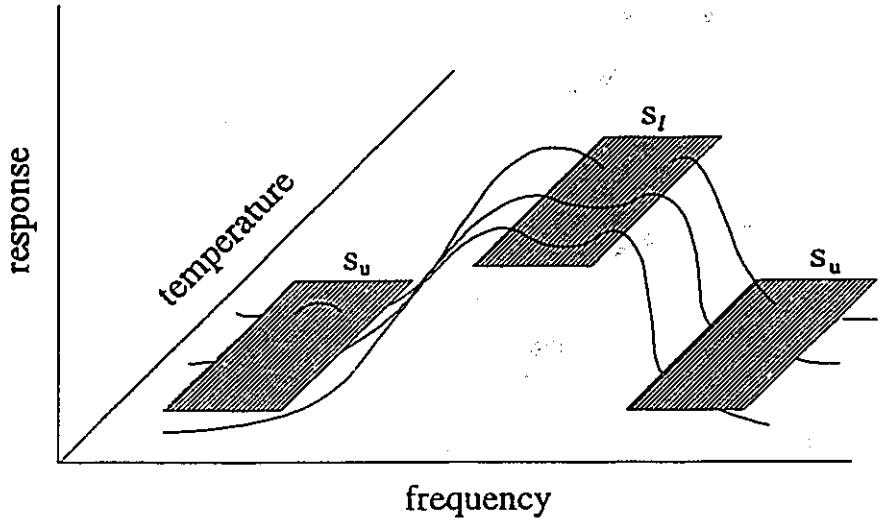

Fig. 2.1 Upper and lower specifications for an amplifier to be designed to operate over a specified temperature range (Bandler and Rizk 1979).

$$e_{ui}(\phi) \triangleq w_{ui}\,(F_m(\phi) - S_{ui}),$$
$$1 \le m \le M, \quad 1 \le i \le L, \tag{2.10}$$

or

$$e_{lj}(\phi) \triangleq w_{lj}(S_{lj} - F_m(\phi)),$$
$$1 \le m \le M, \quad 1 \le j \le L, \tag{2.11}$$

where w_{ui} and w_{lj} are nonnegative weighting factors. Subscripts m and i (or j) may be different, if upper and lower specifications are imposed on the same response. In such a case, S_{ui} and S_{lj} $(i \neq j)$ and $F_m(\phi)$ are the corresponding specifications and response. However, m just indicates its position in (2.5) and is not necessarily equal to either i or j. A positive(non positive) error function implies that the corresponding specification is violated(satisfied).

To unify indices of error functions, we assemble all error functions resulting from (2.10) and (2.11) in a vector

$$e(\phi) = \begin{bmatrix} e_1(\phi) \\ e_2(\phi) \\ \cdot \\ \cdot \\ \cdot \\ e_L(\phi) \end{bmatrix}. \tag{2.12}$$

The acceptable region R_a, with respect to a set of given specifications, is defined by

$$R_a \triangleq \{\phi \mid e(\phi) \le 0\}. \tag{2.13}$$

2.2.4 Yield Estimation

There exist two categories of approaches to yield estimation, namely, the geometrical (or deterministic) method and the Monte Carlo (or statistical) method.

The geometrical method considers a continuous spread of parameter values.

The typical formula to evaluate the yield for the geometrical method is

$$y(\phi^0, \epsilon) = \int_{-\infty}^{\infty} I(\phi) p(\phi, \phi^0, \epsilon) d\phi \ , \tag{2.14}$$

where function $I(\phi)$ is the acceptance index defined by the following

$$I(\phi) = \begin{cases} 1 & \text{if } c(\phi) \leq 0 \\ 0 & \text{otherwise,} \end{cases} \tag{2.15}$$

and $p(\phi, \phi^0, \epsilon)$ is the parameter probability density function.

The Monte Carlo method utilizes a number of sample points generated according to the given probability function of circuit parameters.

$$\begin{aligned} \phi^i &= \phi^0 + \Delta\phi^i, \\ \phi^i &\in R_\epsilon(\phi^0, \epsilon), \quad i = 1, 2, \ldots K, \end{aligned} \tag{2.16}$$

where K is the total number of sample points. These sample points represent circuit outcomes in the manufacturing process. We will refer these sample points as *statistical outcomes*. For the statistical method, the manufacturing yield can be estimated by the following

$$Y(\phi^0, \epsilon) = \frac{1}{K} \sum_{i=1}^{K} I(\phi^i). \tag{2.17}$$

2.2.5 Yield-Driven Circuit Design

An important problem in yield-driven design is design centering (Bandler and Abdel-Malek 1978, Bandler and Kellermann 1983). The purpose of design centering is to enhance yield by optimizing only nominal parameters ϕ^0 and keeping the tolerances fixed. This is a special case of statistical design. The fixed tolerance problem has its practically meaningful application when the manufacturing procedure

always exhibits limited precision as a constant background and model inaccuracy is known. The estimated yield is directly or indirectly used as the design objective. The general formulation to solve such a problem can be stated as

$$\max_{\phi^0} \left\{ y(\phi^0) = \int_{-\infty}^{\infty} I(\phi) p(\phi, \phi^0, \epsilon)\, d\phi \right\} \tag{2.18}$$

for geometrical approaches, or

$$\max_{\phi^0} \left\{ Y(\phi^0) = \frac{1}{K} \sum_{i=1}^{K} I(\phi^i) \right\} \tag{2.19}$$

for Monte Carlo approaches.

For some other cases, it is possible to influence tolerances by adjusting the manufacturing process. Presumably, the smaller the tolerances around a valid nominal design are, the higher the yield is. However, tightening tolerances will result in increased manufacturing cost. Another type of problem involves the design of tolerances. Such a problem is known as optimal tolerancing (Bandler and Abdel-Malek 1978), optimal tolerance assignment (Karafin 1971, 1974), or the variable tolerance problem (Bandler and Kellermann 1983). In such cases, the parameters in ϕ^0, along with those in ϵ, are considered as optimization variables. The objective function consists of costs which reflect assigned tolerances. A yield figure should also be attached as one more design constraint to ensure the desired yield. The typical formulation is

$$\begin{aligned} &\underset{\phi^0, \epsilon}{\text{minimize}}\ C(\phi^0, \epsilon), \\ &\text{subject to} \\ &Y(\phi^0, \epsilon) \geq Y_L, \end{aligned} \tag{2.20}$$

where Y_L is a yield specification. When a 100-percent yield is required, this problem becomes the worst case design problem and the formulation corresponding to (2.20) is

$$\begin{aligned} &\underset{\phi^0,\,\epsilon}{\text{minimize}}\ C(\phi^0, \epsilon), \\ &\text{subject to} \\ &e(\phi) \le 0,\ \forall \phi \in R_\epsilon, \end{aligned} \tag{2.21}$$

where R_ϵ is the tolerance region.

2.3 REVIEW OF APPROACHES TO STATISTICAL DESIGN

The statistical design approach was originated in the early 70's. Among others, Karafin (1971 and 1974), Butler (1971), Pinel and Roberts (1972), Elias (1975), Bandler (1972, 1973 and 1974), Bandler, Liu and Tromp (1976) are pioneers making fundamental contributions to the research. Then, many others, Director and Hachtel (1974, simplicial approximation), Soin and Spence (1978, the center of gravity method), Bandler and Abdel-Malek (1978, updated approximations and cuts), Polak and Sangiovanni-Vincentelli (1979, outer approximation), Tahim and Spence (1979, the radial exploration approach), Antreich and Koblitz (1982, design centering by yield prediction), Styblinski and Ruszczynski (1983, stochastic approximation), Singhal and Pinel (1981, parametric sampling), Bandler and Chen (1988, generalized ℓ_p centering), Biernacki and Styblinski (1986, dynamic constraints approximation), Severson and Simpkins (1987, worst case methods via Hadamard analysis), Purviance and Meehan (1988, sensitivity figure for yield improvement), Vai, Prasad and Meskoob (1990, yield optimization through simulated annealing) have made substantial further contributions. In this section, we describe several representative

approaches.

2.3.1 The Simplicial Approximation Technique

We examine the simplicial approximation technique of Director and Hachtel (1977), and Brayton, Director and Hachtel (1980). Their method is a geometric based one which uses the simpler expressions to approximate the acceptable region.

For a given acceptable region R_a, the boundary of it is defined as

$$\partial R_a = \{ \phi \mid c_i(\phi) \le 0, \forall\, i, \text{ and } c_j(\phi) = 0, \exists\, j,\ i,\ j \in \{1,2,...,L\}\}. \qquad (2.22)$$

Given a set of m ($> n + 1$) points on boundary ∂R_a, a group of (n-1)-dimensional simplices can be obtained. The simplicial approximation to R_a, denoted by $\hat{R}_a$, is a polyhedron whose faces are (n-1)-dimensional simplices. A procedure was suggested by Director and Hachtel (1977) to obtain a convex hull which forms such an approximation to R_a. Specifically, the simplicial approximation is defined by

$$\hat{R}_a = \{ \phi \mid \eta_i^T \phi \le b_i,\ i = 1, 2, ..., N_f \}, \qquad (2.23)$$

where the η_i are outward pointing normals to the bounding hyperplanes defined by n points on ∂R_a, the b_i are the distances between these hyperplanes in the approximation and the origin, and N_f is the total number of hyperplanes. When obtaining a simplicial approximation, we can find an estimate of the new nominal point by determining the center of the largest hypersphere inscribed inside of the polyhedron. A linear programming approach is recommended to find the center of the largest hypersphere.

To improve the simplicial approximation, more and more points on the

boundary should be found to expand the polyhedron. Each of these points is found along the outward normal direction of the largest face extending from the midpoint the same face of the current polyhedron. The nominal point is also updated when a new approximation is available.

A simple illustration of the simplicial approximation is given in Fig. 2.2.

Further development based on the simplicial approximation approach has been made to include arbitrary statistical distributions by Brayton, Director, and Hachtel (1980), and to allow the use of multiple criterion optimization for yield by Lightner and Director (1981). The simplicial approximation approach is based on an essential assumption of a convex acceptable region. This assumption limits its application because determining whether or not a given problem has a convex acceptable region is very difficult by itself.

2.3.2 The Center-of-Gravity Method

Soin and Spence (1980) proposed to use Monte Carlo analysis to constitute a random sampling, or statistical exploration, of the tolerance region. Following Monte Carlo analysis, which identifies each circuit outcome as 'pass' or 'fail' accordingly, the centers of gravity of both the pass and fail outcomes are determined. The center for pass(fail) outcomes, denoted by $G_p(G_f)$, is simply the arithmetic mean of values of the coordinates of pass(fail) circuit outcomes. Then, a new nominal point along the direction from G_f to G_p is found so that this point has a higher yield than the previous nominal point. The proposed relationship between the old and new nominal points is

$$\phi^0_{new} = \phi^0_{old} + \lambda(G_p - G_f). \tag{2.24}$$

The way to calculate the step size, λ, is very decisive in determining the

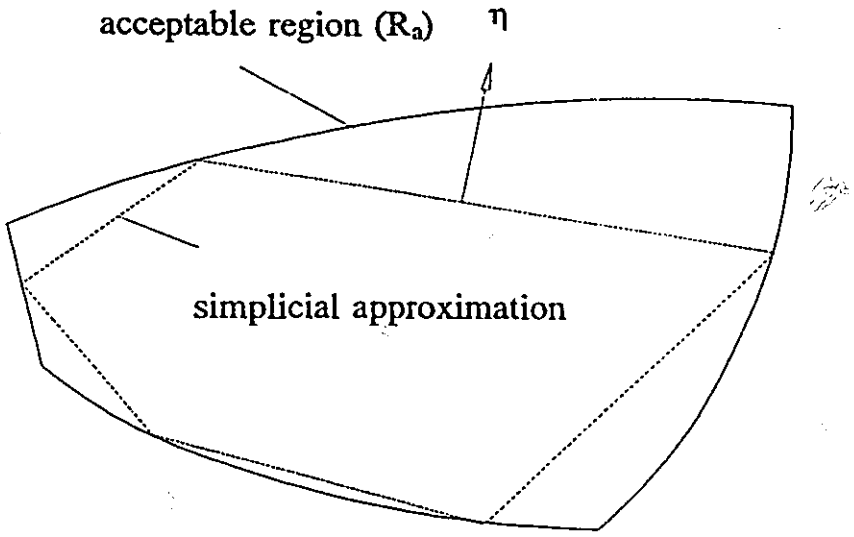

Fig. 2.2 Illustration of the simplicial approximation approach (Director and Hachtel 1977).

computational effort needed for the entire process. A rule for the choice of λ is derived for a very restrictive case where 100% yield can be achieved with the given tolerances and the nominal point with 100% yield must be on the straight line between G_p and G_f. For more general cases, a more precise rule to choose λ is difficult to obtain. Therefore, several λ should be calculated and only the one with highest yield is selected. Many iterations may be necessary before the estimated yield does not increase. Two outcome sampling schemes are developed to achieve a high confidence level at an acceptably low computational cost.

The serious problems with this approach are: (1) the unclear relationship between the two gravity centers and yield, consequently, the uncertainty to find the optimal nominal point along the straight line between two centers; (2) lack of the automatic procedure to find the step size.

2.3.3 Updated Approximations and Cuts

In this method, proposed by Bandler and Abdel-Malek (1978,1980), a lower-order multidimensional polynomial approximation is made to the acceptable region. First, the circuit is simulated at some points, called base points, which are selected according to a derived scheme to preserve one-dimensional convexity/concavity of the circuit response function. Then, the coefficients of the approximation function, a quadratic polynomial in their case, can be determined by solving a set of linear equations based on the simulated response values and the coordinates of the base points. To prevent possible loss of accuracy at some points considered as very critical in worst case design, multiple regions around these points are suggested be approximated individually. It is also necessary to update the approximation functions

as the nominal point moves.

An analytical approach to the evaluation of yield and yield sensitivities is proposed based on computation of the hypervolume of the intersection of the non-acceptable region and the tolerance region. The hypervolume formula is derived from linear cuts of the tolerance region. The linear cuts are functions of the nonlinear constraints defining the boundary of the acceptable region. The scheme to construct the linear cut from the linear or quadratic constraints, which approximate the original expensive nonlinear constraints, is given. Sensitivities of the yield estimate are also available to be used in a gradient-based optimization algorithm.

2.3.4 Stochastic Approximation

Styblinski and Ruszczynski (1983) found an analogy between statistical design centering and the problem of finding the maximum of the regression function

$$\max_{\phi^0} \left\{ Y(\phi^0) = \int_{-\infty}^{\infty} I(\phi)\, p(\theta)\, d\theta = E\{I(\phi)\} \right\}, \tag{2.25}$$

where θ is an n-dimensional random variable with zero expectation, the outcome $\phi = \phi^0 - \theta$, and $p(\theta)$ is the probability density function of θ. It is evident that (2.25) can be obtained by redefining the origin of the parameter space.

The stochastic approximation approach, then, is used to solve the problem in (2.25). The following iterations are used

$$\begin{aligned} \phi^0_{(k+1)} &= \phi^0_k + \tau_k d_k \\ d_k &= (1 - \rho_k) d_{(k-1)} + \rho_k \xi_k, \qquad 0 \le \rho_k \le 1 \end{aligned} \tag{2.26}$$

where k is the current iteration, ξ_k an estimate of the gradient of $Y(\phi^0)$, and $\tau_k \to 0$ and $\rho_k \to 0$ are non-negative coefficients. The significant advantage of this approach

is that the gradient of yield, ξ_k, can be estimated based on only one outcome point. Formulas to generate τ_k and ρ_k are given.

From the example given by the authors, a dramatic yield increase is observed during the first several iterations of the optimization process. It is suspected that the selected starting nominal point out of the acceptable region might contribute to the fast initial convergence. The algorithm exhibits slow convergence when close to the solution.

Like many stochastic algorithms, this method may suffer from very slow convergence rate when approaching the solution. Another limitation of this method is that the gradient evaluation restricted to differentiable functions.

2.3.5 Parametric Sampling

The key principle of the approach due to Singhal and Pinel (1981) is to replace the original probability density function by some other density function

$$y(\phi^0) = \int_{-\infty}^{\infty} \left[I(\phi) \frac{p(\phi, \phi^0)}{h(\phi, \phi^0)} \right] h(\phi, \phi^0) d\phi \ , \tag{2.27}$$

where $h(\phi, \phi^0)$ is a chosen sampling density. $h(\phi, \phi^0)$ can be arbitrary except for the requirement that $h(\phi, \phi^0) \neq 0$ whenever $I(\phi)p(\phi, \phi^0) \neq 0$. The corresponding Monte Carlo version of (2.27) is

$$y(\phi^0) = \frac{1}{K} \sum_{i=1}^{K} \left[I(\phi^i) \frac{p(\phi^i)}{h(\phi^i)} \right] = \frac{1}{K} \sum_{i=1}^{K} I(\phi^i) W(\phi^i) \tag{2.28}$$

where ϕ^i are sample points generated from the sampling density $h(\phi^i)$ rather than the component density $p(\phi^i)$. The weight factors

$$W(\phi^i) = \frac{p(\phi^i)}{h(\phi^i)}$$

compensate for the use of a different density function.

The Monte Carlo analysis is carried out once only using a sampling distribution $h(\phi^i)$. The sampled statistical outcomes and the circuit responses are stored in a database. The advantage of the parametric sampling approach is that no new circuit simulations are required during the optimization if the tolerance region does not move out of the territory of the database. However, when the optimization leads some outcomes out of the database, a dynamic updating scheme may be necessary to enlarge the database.

2.3.6 Sensitivity Figure for Yield Improvement

Purviance and Meehan (1988) introduced a sensitivity figure for use in gradient optimization. The main hypothesis of their approach is that use of this new sensitivity figure in a gradient-based optimization process will result in a circuit design with improved yield.

Suppose that the circuit performance of interest is $F(\phi)$. The sensitivity to a parameter ϕ_j is calculated as (Gupta, Garg and Chadha 1981)

$$T(F, \phi_j) = \frac{\phi_j}{|F|} \frac{\partial |F|}{\partial \phi_j} \Big|_{\phi=\phi^0} . \tag{2.29}$$

The response function is first normalized with respect to the nominal point and the tolerance extremes such that the normalized nominal point is at the origin and the normalized tolerance extremes are -1 and 1. Then the function is expanded as a polynomial function at the nominal point

$$F(\phi) \approx g(\bar{\phi}) \triangleq g_0 + \sum_{i=1}^{N} \alpha_i \bar{\phi}_i + \sum_{i,j=1}^{N} \alpha_i \alpha_j \bar{\phi}_i \bar{\phi}_j + \sum_{i,j,k=1}^{N} \alpha_i \alpha_j \alpha_k \bar{\phi}_i \bar{\phi}_j \bar{\phi}_k + \ldots, \tag{2.30}$$

where $\bar{\phi}_j$ stands for the normalized variable. Define $\overline{g}(\bar{\phi}_j)$ as the average value of $g(\bar{\phi})$ with respect to all the parameters except $\bar{\phi}_i$, that is

$$\overline{g}(\bar{\phi}_i) = \int_{-1}^{1} g(\bar{\phi}) p(\bar{\phi}_1, \ldots, \bar{\phi}_{i-1}, \bar{\phi}_{i+1}, \ldots, \bar{\phi}_N) d\bar{\phi}_1 \cdots d\bar{\phi}_{i-1} d\bar{\phi}_{i+1} \cdots d\bar{\phi}_N , \tag{2.31}$$

where $p(\bar{\phi}_1, \ldots, \bar{\phi}_{i-1}, \bar{\phi}_{i+1}, \ldots, \bar{\phi}_N)$ is the probability density function of all the normalized parameters except $\bar{\phi}_i$. Notice that an assumption made is that it is possible to separate $\bar{\phi}_i$ from the probability density function of all parameters. The derivative of the average performance function value is given by

$$\frac{\partial \overline{g}(\bar{\phi}_i)}{\partial \bar{\phi}_i} = \int_{-\infty}^{\infty} \frac{\partial g(\bar{\phi})}{\partial \bar{\phi}_i} p(\bar{\phi}_1, \ldots, \bar{\phi}_{i-1}, \bar{\phi}_{i+1}, \ldots, \bar{\phi}_N) d\bar{\phi}_1 \cdots d\bar{\phi}_{i-1} d\bar{\phi}_{i+1} \cdots d\bar{\phi}_N \tag{2.32}$$

The derivatives

$$\frac{\partial \overline{g}(\bar{\phi}_i)}{\partial \bar{\phi}_i}, \; i = 1, 2, \ldots, N,$$

are used as the gradient required by the optimizer to influence the optimization solution.

The problems with this approach are (1) the assumption on the separable probability density function might not be applicable to many practical problems, and (2) the unclear connection between the solution found by the proposed sensitivity figure and the nominal point for optimal yield.

2.3.7 Simulated Annealing Optimization

Simulated annealing is a technique for the solution of difficult combinatorial optimization problems (Kirkpatrick, Gelatt and Vecchi 1983). It has been extensively used for circuit geometrical partitioning, component placement, and circuit wiring (Jepsen and Gelatt 1983, Casotto, Romeo, and Sangiovanni-Vincentelli 1987). Very recently, this method was applied to modelling of microwave semiconductor devices (Vai, Prasad, Li, and Kai 1989). Vai, Prasad, and Meskoob (1990) also used simulated annealing optimization for microwave circuit yield-driven design.

Simulated annealing belongs to the random optimization category. It conditionally accepts high intermediate values of the objective function to allow probabilistic hill-climbing. Following the mechanism of the annealing process, a controlling parameter called pseudo-temperature is used in the optimization procedure. Pseudo-temperature is set relatively high at the initial stage, then is decreased artificially and slowly. If the present iterative solution decreases the objective function value, the solution is accepted as in conventional optimization methods. If an intermediate solution increases the objective function value, then the acceptance is conditional on the result of a random experiment such that probability of the acceptance obeys a Boltzmann distribution

$$e^{\frac{-\alpha \Delta U_k}{T_k}} \tag{2.33}$$

where ΔU_k is the difference between two objective function values, T_k is the current temperature, and α is a weighting factor. For the same amount of objective function increment, it is more likely to be accepted at a higher temperature than at a lower one.

This optimization procedure is applied to a distributed amplifier circuit (Vai,

Prasad, and Meskoob 1990). The objective function for yield optimization is defined using the predicted yield rate. (Unfortunately, the formulation of the objective function for yield optimization is not made available in their paper). Solutions with improved yield rates are accepted automatically while the acceptance of those with decreased yield rates is governed by the Boltzmann distribution. A interesting advantage of their approach is that they allow the number of stages in the distributed amplifier as variable.

Simulated annealing optimization requires extremely little computational effort itself and is very easy to implement. The most distinct advantage of simulated annealing is its ability to reach the global optimal solution without requiring a good initial starting point. However, it is commonly admitted that yield optimization should start with a nominal design which usually is a reasonably good point. The major drawback of simulated annealing optimization is the very slow convergence rate when the temperature is low. Very high computational costs due to repeated circuit simulation may need be required.

2.4 THE GENERALIZED ℓ_p CENTERING APPROACH

2.4.1 Formulation of Yield Optimization

It is always our desire to convert the problem of yield optimization to a well behaved mathematical programming problem so that modern mathematical optimization techniques can be applied. Bandler and Chen (1989) proposed the generalized ℓ_p centering approach.

In the following the design variables are the nominal values ϕ^0. K statistical outcomes are generated from the given probability density function. Although only

the outcomes ϕ^i appear in the error functions, they depend on ϕ^0 because the ϕ^i are related to ϕ^0.

After the error vector e^i for the outcome ϕ^i has been assembled as

$$e^i = \begin{bmatrix} e_1(\phi^i) \\ e_1(\phi^i) \\ \cdot \\ \cdot \\ \cdot \\ e_L(\phi^i) \end{bmatrix} \tag{2.34}$$

where L is the total number of errors considered, the formulation of the objective function for optimization can follow the procedure described in Bandler and Chen (1988). First, we create the generalized ℓ_p function v^i from e^i,

$$v^i = \begin{cases} \left[\sum_{j \in J(\phi^i)} (e_j(\phi^i))^p \right]^{\frac{1}{p}}, & \text{if } J(\phi^i) \neq \varnothing \\ -\left[\sum_{j=1}^{L} (-e_j(\phi^i))^{-p} \right]^{-\frac{1}{p}}, & \text{if } J(\phi^i) = \varnothing \end{cases} \tag{2.35}$$

where

$$J(\phi^i) = \{ j | e_j(\phi^i) \geq 0 \}. \tag{2.36}$$

Then we define the one-sided ℓ_1 objective function for yield optimization as

$$U(\phi^0) = \sum_{i \in I} \alpha_i v^i , \tag{2.37}$$

where

$$I = \{ i | v^i > 0 \} \tag{2.38}$$

and α_i are positive multipliers. If the α_i were chosen as

$$\alpha_i = \frac{1}{|v^i|} \tag{2.39}$$

then function $U(\phi^0)$ would become the exact number of unacceptable circuits, that is,

$$U(\phi^0) = \frac{\text{number of unacceptable circuits}}{K} \tag{2.40}$$

and the yield would be

$$Y(\phi^0) = 1 - \frac{U(\phi^0)}{K}. \tag{2.41}$$

The mechanism of the one-sided ℓ_1 function naturally imitates the relation between the yield and unacceptable or acceptable outcomes. Now, the task of maximizing yield Y is converted to one of minimizing $U(\phi^0)$. That is

$$\underset{\phi^0}{\text{minimize}}\, U(\phi^0) \ . \tag{2.42}$$

We use (2.39) to assign multipliers α_i at the starting point and fix them during the optimization process. Then $U(\phi^0)$ is no longer the count of unacceptable outcomes during optimization, but a continuous approximate function to it.

2.4.2 Implementational Aspects

Suppose the value of p in (2.35) is chosen as 1. The objective function for the one-sided ℓ_1 optimization becomes

$$U(\phi^0) = \sum_{i \in I} \sum_{j \in J(\phi^i)} \alpha_i e_j(\phi^i) \tag{2.43}$$

where α_i, I and $J(\phi^i)$ are defined as before. In (2.43), error functions for optimization are $e_j(\phi^i)$. In (2.37), error functions for optimization are v^i. The functions $e_j(\phi^i)$ are differentiable, but the functions v^i of (2.35) may not be.

However, (2.43) has more error functions than (2.37) if more than one specification is imposed.

In the following chapters, we use both (2.37) and (2.43) in the different yield optimization problems. To distinguish them, we refer to (2.37) and (2.43) as Implementation I and Implementation II of the one-sided ℓ_1 centering approach, respectively.

Several reoptimizations with updated α_i may be applied to further increase yield. Each can use a different number of statistical outcomes or a different set of outcomes.

To summarize the discussion in this section, all steps involved in our yield optimization are shown in Fig. 2.3.

2.5 ONE-SIDED ℓ_1 MATHEMATICAL PROGRAMMING

A highly efficient one-sided ℓ_1 optimization algorithm (Bandler, Chen and Madsen 1988) is used to solve (2.37) or (2.43). The algorithm is based on a two-stage method combining a first-order method, the trust region Gauss-Newton method, with a second-order method, the quasi-Newton method. Switching between the two methods is automatically made to ensure global convergence of the combined algorithm.

2.5.1 Formulation of the Problem

The optimization problem to be considered has the following mathematical formulation. Let

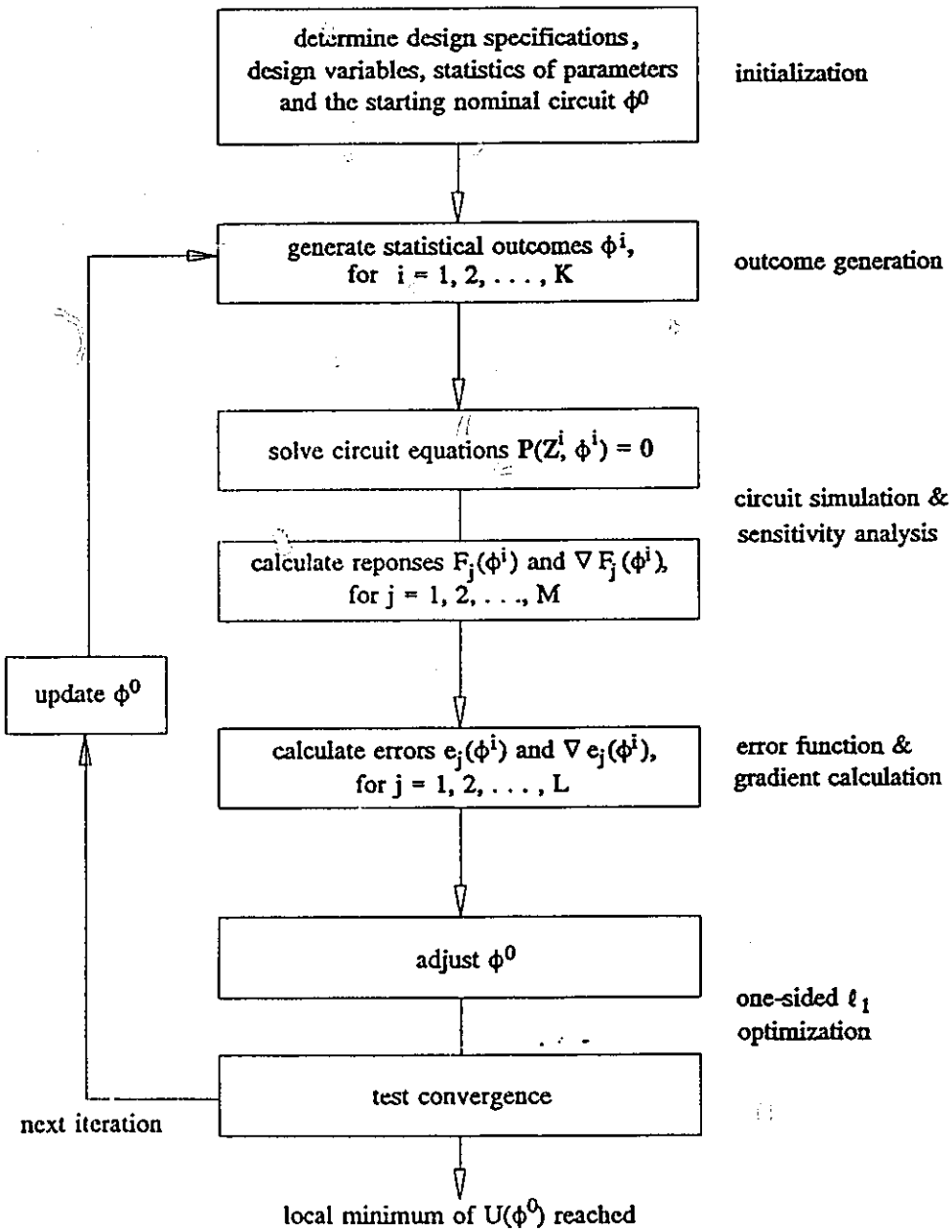

Fig. 2.3 Flowchart of yield optimization

$$f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \cdot \\ \cdot \\ \cdot \\ f_m(x) \end{bmatrix} \tag{2.44}$$

be a set of nonlinear, continuously differentiable functions. The vector

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_N \end{bmatrix} \tag{2.45}$$

is the set of parameters to be optimized. We define an index set which contains indices of all functions with positive values

$$J(x) = \{ j \mid f_j(x) > 0 \}. \tag{2.46}$$

Then the one-sided ℓ_1 optimization problem can be stated as

$$\underset{x}{\text{minimize}} \left\{ U(x) = \sum_{j \in J(x)} f_j(x) \right\}. \tag{2.47}$$

Substituting either (2.37) or (2.43) into (2.42) gives a one-sided ℓ_1 problem as defined in (2.47).

2.5.2 Algorithms for the One-Sided ℓ_1 Problem

The trust region Gauss-Newton and Quasi-Newton methods have their own advantages. The trust region method is supposed to work well at the beginning stage of optimization, but to converge very slowly when close to a singular solution. The quasi-Newton method has a fast rate of convergence near a solution but is not reliable

from a bad starting point. By combining them together, we hope to fully explore their advantages and to effectively avoid their shortcomings (Hald and Madsen 1981, 1985). Several switches between the two methods may take place and the switching criteria ensure global convergence of the combined algorithm.

A Trust Region Gauss-Newton Method (Method I)

At the kth iteration with the present solution x_k, a local bound Λ_k is chosen. The following subproblem is to be solved

$$\begin{aligned} &\underset{h,y}{\text{minimize}} \sum_{j=1}^{m} y_j \\ &\text{subject to} \\ &y_j \geq f_j(x_k) + f_j'(x_k)^T h, \quad j = 1, 2, \ldots, m, \\ &y_j \geq 0, \\ &\Lambda_k \geq h_i, \quad \Lambda_k \geq -h_i, \quad i = 1, 2, \ldots, N, \end{aligned} \tag{2.48}$$

where f_j' is the gradient of f_j. A standard linear programming algorithm can be used to solve this problem for h_k. If h_k reduces the objective function of the original problem, i.e., if $U(x_k + h_k) < U(x_k)$, then $x_k + h_k$ is accepted as an improved solution to (2.48) and the next iterate starts. Otherwise, this iteration is considered as a failure, and the previous solution is kept.

The local bound Λ_k should be adjusted in every iteration. The rule for adjusting Λ_k is based on whether or not the linearized subproblem (2.48) is a good approximation to the original nonlinear problem in the present trust region defined by the current local bound Λ_k. A detailed description of the algorithm to control Λ_k is given by Bandler, Kellermann and Madsen (1987).

A Quasi-Newton Method (Method II)

This method solves the optimality condition of the one-sided ℓ_1 problem. The optimality condition is a set of equations

$$\sum_{j \in J_+} f_j'(x) + \sum_{j \in J_0} \delta_j f_j'(x) = 0, \tag{2.49}$$

where J_+ and J_0 are defined as

$$J_+ \triangleq \{ j \mid f_j(x) > 0 \} \tag{2.50}$$

and

$$J_0 \triangleq \{ j \mid f_j(x) = 0 \}, \tag{2.51}$$

respectively, and the multipliers must satisfy

$$1 \geq \delta_j \geq 0, \qquad j \in J_0. \tag{2.52}$$

These optimality equations result from applying the Kuhn-Tucker conditions to the one-sided ℓ_1 problem.

The quasi-Newton method is used to solve (2.49). Second-order derivative information is required. A modified BFGS formula (Powell 1978, and Bandler, Kellermann and Madsen 1987) can be adopted to generate and update the Hessian matrix.

A Combined 2-Stage Algorithm

Based on the theory of Hald and Madsen (1981 and 1985), the algorithm combines the trust region Gauss-Newton method (Method I) with the quasi-Newton method (Method II).

At the beginning, the trust region is used. Additional to the computation of the trust region method, the preparation for Method II is also made, which includes

estimation of active sets J_0 in (2.51), estimation of the multipliers δ_i in (2.52), and updating of the approximate Jacobian using a modified BFGS formula (Powell 1978, also Bandler, Kellermann and Madsen 1987).

A switch from Method I to Method II is made if the following conditions are met:

(1) The estimated active set J_0 has been unchanged over a predetermined number of consecutive iterations.

(2) The estimated multipliers corresponding to J_0 satisfy (2.52).

If Method II is unsuccessful, then the safe method, Method I, should be used again. A switch from Method II back to Method I is made if one of the following conditions is met:

(1) The contents of either J_0 or J_+ need be updated because a function not included in J_0 has become zero or changed sign.

(2) At least one multiplier has violated the constraint.

(3) A quasi-Newton step fails to decrease the residual of the optimality equations:

$$\|R_{k+1}\| < 0.999\,\|R_k\| \tag{2.53}$$

Several switches between the two methods may take place until convergence is reached. This two-stage algorithm will be used in our yield optimization in the following chapters.

2.6 CONCLUDING REMARKS

In this chapter, we have considered the yield-driven circuit design problem. The relevant definitions and notations have been introduced. The formulation of the statistical design problem has been presented. A number of existing approaches to

statistical design have been reviewed in certain detail. Their individual advantages and shortcomings have been addressed.

In order to introduce the new results presented in the upcoming chapters, we have emphasised on the generalized ℓ_p centering approach. Due to its important properties, a special case of this approach, namely, the one-sided ℓ_1 approach, has been discussed. Two possible implementations of the one-sided ℓ_1 method have been identified. A two-stage algorithm combining the trust Gauss-Newton method and the quasi-Newton method has been outlined.

3

EFFICIENT QUADRATIC APPROXIMATION FOR STATISTICAL DESIGN

3.1 INTRODUCTION

In order to make existing statistical circuit design methods more practically usable, many approaches have been devised to reduce very costly computational effort by approximating acceptable regions or circuit responses. Quadratic approximation has proven suitable and successful (Bandler and Abdel-Malek 1978, Abdel-Malek and Bandler 1980, Hocevar, Lightner and Trick 1983, and Biernacki and Styblinski 1986). However, the determination of a quadratic model itself for a problem with a large number of variables may be too expensive.

For a circuit with 50 elements, the number of coefficients in the quadratic model is 1326. The calculation of the coefficients in a traditional manner involves 1326 circuit simulations and solving a linear system of 1326 equations. Besides all the coefficients, the matrix of the linear system requires storage of a 1326 by 1326 array. Determining a quadratic approximation to the response of such a circuit creates quite a large problem in terms of computer time and storage, although the circuit itself may be of a moderate scale. Therefore, for large scale problems the traditional approaches that aim to obtain unique quadratic models do not effectively reduce computational costs.

Biernacki and Styblinski (1986) introduced the concept of the maximally flat interpolation and presented an updating algorithm. The most significant property of their approach is that the method allows the number of actual circuit simulations

required for an accurate model to be much less than that needed for a full unique quadratic approximation. However, the computational requirement of the method, especially storage space, is still high.

In this chapter we substantially enhance the maximally flat quadratic interpolation. Our approach makes use of a fixed pattern of points at which simulation is performed, resulting in very low computational requirements for both CPU time and storage (Biernacki, Bandler, Song and Zhang 1989). The basic concept is reviewed in Section 3.2. Our new approach is described in Section 3.3. Comparisons of the efficiency of our approach and that of the original maximally flat quadratic approximation are given in the same section. In Section 3.4, the proposed quadratic approximation technique is applied to model circuit response functions. A low-pass filter serves as an example to demonstrate this implementation. In Section 3.5, we utilize the quadratic approximation to model not only circuit performance functions, but also their gradients (Bandler, Biernacki, Chen, Song, Ye and Zhang 1991c). Our gradient-based optimization procedure, the one-sided ℓ_1 centering approach (Bandler and Chen 1989), requires gradient information. Higher gradient accuracy will improve the overall performance of the optimization process.

Finally, Section 3.6 contains the conclusions.

3.2 THE MAXIMALLY FLAT QUADRATIC APPROXIMATION

A quadratic model in polynomial form to be used to approximate a given function $f(x)$, $x = \left[x_1 \;\; x_2 \;\; \ldots \;\; x_n\right]^T$, can be written as

$$q(x) = a_0 + \sum_{i=1}^{n} a_i(x_i - r_i) + \sum_{i,j=1,i\le j}^{n} a_{ij}(x_i - r_i)(x_j - r_j), \tag{3.1}$$

where $r = [r_1 \ r_2 \ . \ . \ . \ r_n]^T$ is a known reference point. The form of the quadratic function used is similar to that of Biernacki and Styblinski (1986). However, $q(x)$ is defined here w.r.t. the reference point r rather than w.r.t. the origin. Note that the subscript notation is such that each coefficient can be easily identified with its corresponding x term, e.g., a_{ij} is the coefficient of $(x_i - r_i)(x_j - r_j)$. Determining a quadratic model is equivalent to determining all its coefficients, which are now unknowns in (3.1).

Suppose that m ($m > n + 1$) evaluations of $f(x)$ are performed at some points x^i, $i = 1, 2, \ldots, m$. These points are called the base points. Using the values of $f(x^i)$, we set up a system of linear equations

$$\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} a \\ v \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, \tag{3.2}$$

where a and v are arranged to have the following orders:

$$a = [\, a_0 \ \ a_1 \ \ . \ \ . \ \ . \ \ a_n \,]^T \tag{3.3}$$

and

$$v = [\, a_{11} \ \ a_{22} \ \ . \ \ . \ \ . \ \ a_{nn} \ \ a_{12} \ \ a_{13} \ \ . \ \ . \ \ . \ \ a_{n-1,n} \,]^T, \tag{3.4}$$

respectively. The vectors f_1 and f_1 are of dimensions $(n + 1)$ and $m - (n + 1)$, respectively. They contain function values $f(x^i)$. The matrix Q_{ij}, $i, j = 1, 2$, is determined by the coordinates of the base points and of r.

Similarly to the approach due to Biernacki and Styblinski (1986), the reduced system with variables v is obtained as

$$C v = e, \tag{3.5}$$

where

$$C = Q_{22} - Q_{21}Q_{11}^{-1}Q_{12} \tag{3.6}$$

and

$$e = f_2 - Q_{21}Q_{11}^{-1}f_1. \tag{3.7}$$

If $m < \frac{(n+1)(n+2)}{2}$, the above system is under-determined.

When the least-squares constraint is applied to v, the unique solution to (3.5) can be found as

$$v = C^T(CC^T)^{-1}e \tag{3.8}$$

and a is readily obtained as

$$a = Q_{11}^{-1}f_1 - Q_{11}^{-1}Q_{12}v. \tag{3.9}$$

Then v, called the minimal Euclidean norm solution of (3.5), and a give the maximally flat quadratic interpolation in the form of (3.1) to $f(x)$. The term of the maximally flat quadratic approximation comes from the mechanism of the least-squares constraint that forces the second-order derivatives to be as small as possible.

3.3 APPROACH USING A FIXED PATTERN OF BASE POINTS

3.3.1 Derivation and Algorithm of the Approach

In the original scheme of Biernacki and Styblinski (1986) all base points are randomly selected. This type of selection allows certain freedom. However, several large matrices have to be stored and manipulated. For instance, matrix C in (3.6) needs an array with dimension $\frac{n(n+1)[m-(n+1)]}{2}$. Meanwhile, some fairly involved calculations, such as matrix inversion, or equivalent calculations shown in (3.6)-(3.9), are required. Even a circuit of a reasonable size may demand large storage space and CPU time. Here, we shall propose a new approach which is based on a

fixed pattern of base points. The regularity of the pattern will greatly reduce storage and simplify the calculation of coefficients.

In our approach, only m ($n + 1 < m < 2n + 1$) base points are used. The reference point r is selected as the first base point x^1. The next n base points are determined by perturbing one variable at a time around r, i.e.,

$$x^{i+1} = r + \left[0 \ . \ . \ . \ 0 \ \beta_i \ 0 \ . \ . \ . \ 0\right]^T, \qquad i = 1, 2, \ldots, n, \tag{3.10}$$

where β_i is a predetermined perturbation. It can be shown that the first ($n + 1$) base points lead to very simple forms of matrices Q_{11}^{-1} and $Q_{11}^{-1}Q_{12}$. They are

$$Q_{11}^{-1} = \begin{bmatrix} 1 & 0 & & & 0 \\ -\frac{1}{\beta_1} & \frac{1}{\beta_1} & . & . & 0 \\ -\frac{1}{\beta_2} & 0 & & & 0 \\ . & & . & & 0 \\ . & 0 & & . & \\ -\frac{1}{\beta_n} & 0 & . & . & \frac{1}{\beta_n} \end{bmatrix} \tag{3.11}$$

and

$$Q_{11}^{-1}Q_{12} = \left[\begin{array}{cccccc|c} 0 & 0 & . & . & . & 0 & \\ \beta_1 & 0 & . & . & . & 0 & \\ 0 & \beta_2 & . & . & . & 0 & \\ 0 & 0 & . & & & 0 & 0 \\ 0 & 0 & & . & & 0 & \\ 0 & 0 & & & . & 0 & \\ 0 & 0 & & & & \beta_n & \end{array}\right]. \tag{3.12}$$

Because of this simple pattern they need not be stored in matrix form.

After the first $n + 1$ base points, the remaining $m - (n + 1)$ points follow to

provide the second-order information on the function. Similarly to the base points defined in (3.10), consecutive base points are selected by also perturbing one variable at a time. For simplicity, these base points are determined by consecutively perturbing the variables in r, that is

$$x^{n+1+i} = r + [0 \ . \ . \ . \ 0 \ \gamma_i \ 0 \ . \ . \ . \ 0]^T, \qquad i = 1, 2, \ldots, k, \tag{3.13}$$

where γ_i is another perturbation of r, which must not equal β_i, and

$$k = m - (n + 1).$$

Under this arrangement matrices Q_{21} and Q_{22} have regular structures. Substituting Q_{21} and Q_{22} into (3.6), the matrix C takes a concise analytical form

$$C = \begin{bmatrix} (\gamma_1 - \beta_i)\gamma_i & & & 0 \\ & \ddots & & \vdots \\ & & 0 & 0 \\ & & & 0 \\ & & & 0 \\ & & (\gamma_i - \beta_i)\gamma_i & 0 \\ & & & 0 \\ & & & 0 \\ & 0 & & 0 \\ & & (\gamma_k - \beta_k)\gamma_k & 0 \end{bmatrix} \tag{3.14}$$

and the vector e can be expressed by

$$e = f_2 - \begin{bmatrix} 1 - \frac{\gamma_1}{\beta_1} & \frac{\gamma_1}{\beta_1} & & & & 0 \\ 1 - \frac{\gamma_2}{\beta_2} & & \frac{\gamma_2}{\beta_2} & & 0 & 0 \\ & & & \cdot & & 0 \\ & & & & \cdot & 0 \\ 1 - \frac{\gamma_i}{\beta_i} & & & \frac{\gamma_i}{\beta_i} & & 0 \\ & & & & \cdot & 0 \\ & & 0 & & \cdot & 0 \\ 1 - \frac{\gamma_k}{\beta_k} & & & & \frac{\gamma_k}{\beta_k} & 0 \end{bmatrix} f_1 . \qquad (3.15)$$

Substituting (3.14) and (3.15) into (3.8), the coefficients are determined by

$$a_{ii} = \frac{1}{\gamma_i - \beta_i} \left[\frac{f(x^{n+1+i}) - f(x^1)}{\gamma_i} - \frac{f(x^{i+1}) - f(x^1)}{\beta_i} \right], \qquad (3.16a)$$

$$i = 1, 2, \ldots, k,$$

$$a_{ii} = 0, \quad i = k+1, \ldots, n, \qquad (3.16b)$$

and

$$a_{ij} = 0, \quad i \neq j, \quad i, j = 1, 2, \ldots, n. \qquad (3.16c)$$

The coefficients a_0 and a_i are easily obtained as

$$a_0 = f(x^1) \qquad (3.17a)$$

and

$$a_i = \frac{[f(x^{i+1}) - f(x^1)]}{\beta_i} - \beta_i a_{ii}, \quad i = 1, 2, \ldots, n. \qquad (3.17b)$$

The maximally flat quadratic interpolation, using a fixed pattern of base points defined here, has an interesting property. All the coefficients of the mixed terms, a_{ij} for $i \neq j$, are conveniently forced to be zero because no related information can be extracted from the fixed pattern. Any of the a_{ii}'s in (3.16a), $i \leq k$, can be

nonzero because double perturbations are made along a straight line parallel to the ith axis. If a third perturbation is made along the same line, it can be shown that the C matrix will not have full row rank, and, therefore, the third perturbation does not provide any extra useful information for the quadratic model. It should be noted, however, that the fact that the mixed term coefficients become zero is a consequence of the maximally flat interpolation.

In a situation where the mixed terms are important, this approach can easily be modified by introducing an appropriate transformation of variables. In such a case the perturbations can be carried out along lines not necessarily parallel to the axes. The proposed fixed pattern of base points can thus be generalized while preserving the main advantages of our approach.

Theoretically speaking, the efficiency and simplicity of our models are achieved at the expense of some model accuracy. It should be stressed, however, that even so-called exact circuit simulation carries certain approximations of actual physical behaviour. Therefore, our approach provides an excellent modeling technique for many practical problems, especially when a very high accuracy is not really necessary. The method is suitable for a maximum number of base points of $2n+1$. It takes advantage of the concept of maximally flat quadratic interpolation and thus any number of base points between $n+1$ and $2n+1$ can be used.

3.3.2 Computational Efficiency

A dynamic updating scheme was proposed by Biernacki and Styblinski (1986), which allows the existing model to be revised when a new base point is added. However, this simple updating may not be suitable if some of the base points are far

from the region of interest. To maintain accuracy, it may be desirable to disregard such points. Our method can be used to rebuild the model very efficiently whenever it is needed. However, it can also be used within the concept of dynamic updating provided that the base points are selected in the aforementioned manner and their number does not exceed $2n+1$.

In this section we compare computational efficiency of our approach with that of the original method of Biernacki and Styblinski (1986). To unify the comparison we assume that exactly $2n+1$ base points are used to build the model. In our approach the required storage is reduced to a minimum. Only $2n$ perturbations and $2n+1$ function values are to be stored. No matrix manipulations are needed. All calculations are simplified to (3.16) and (3.17). The operational count to calculate all coefficients using this pattern can be merely $4n$. In this new approach, the storage requirement and computational count vary linearly with the number of variables. For the original method of Biernacki and Styblinski (1986), at least, all base points, matrices Q_{11} and C are stored in three arrays with dimensions $n\times(2n+1)$, $(n+1)\times(n+1)$ and $n\times n\times(n+1)/2$, respectively, and the computational count is $O(n^4)$. For the original approach, the storage requirement and computational count vary cubically and quarticly, respectively, with the number of variables. For a circuit with 50 elements and m chosen as 101, we need storage consisting of two arrays of 101 and 100, and computational effort of 200 multiplications. The original approach would require storage consisting of three arrays of 50 by 101, 51 by 51 and 50 by 1275, respectively, and incomparable computational effort.

Statistical design or design centering involves a very large number of circuit simulations. It is essential to reduce the CPU time due to simulation which usually

takes a large portion of overall design time. In such design, the trend of the circuit response hypersurface is much more important than the accuracy of the individual circuit responses. Therefore, our maximally flat quadratic approximation is capable of accelerating the design process without loosing much accuracy in the vicinity of a nominal circuit. The approximate model is evaluated at all design outcomes that are sampled according to the characterized statistical properties of the variables, such as means, variances and dependencies.

3.4 QUADRATIC APPROXIMATION OF CIRCUIT RESPONSES

3.4.1 Implementation

The quadratic approximation technique has been implemented within the framework of a circuit design program which uses a general-purpose simulator and implementation I of the generalized ℓ_1 centering approach outlined in the preceding section. The individual circuit responses are approximated by quadratic functions. The reference point is defined as the nominal point. At each iteration, a set of quadratic models is built. The models are evaluated for all outcomes, i.e, the statistically sampled circuits. The objective function is calculated from the resulting approximate error functions.

3.4.2 Design of a 11-Element Low-Pass Filter

A low-pass ladder filter with 11 elements used by Singhal and Pinel 1981, and Wehrhahn and Spence 1984, shown in Fig.3.1, was used in this example. The upper specification was 0.32dB in the frequency range from 0.02Hz to 1Hz, and the lower specification 52dB at 1.3Hz, on the insertion loss. The frequency sample points are

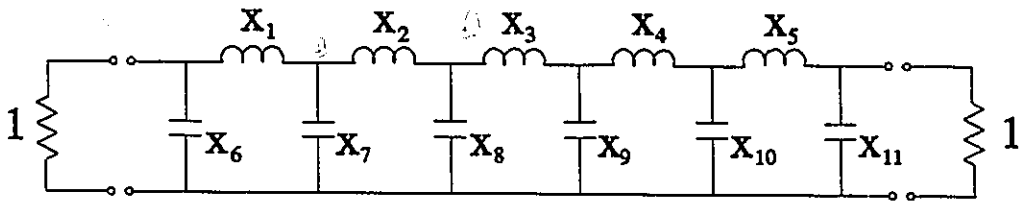

Fig. 3.1 The LC low-pass filter.

(0.02, 0.04, ..., 1, 1.3Hz). A relative tolerance of 1.5% was assumed for all elements. Outcomes were uniformly distributed between tolerance extremes. The starting point was the result of a synthesis procedure (Singhal and Pinel 1981).

Results and comparisons are given in Table 3.1. Implementation I of the generalized ℓ_1 centering approach was used. Two designs within the same optimization environment were carried out. The only difference between the two approaches was the way of calculating the circuit responses. Actual circuit simulations were used in the first design and our approximation method was utilized in the second one. Each design consisted of two successive centering processes shown as phase 1 and phase 2 in Table 3.1. Two phases of design with actual simulations took approximately 158.6 minutes on the VAX 8600 and required 79200 circuit simulations. The final yield was 63.7%. The very large number of circuit simulations was due to repeated circuit simulation of 200 outcomes, gradient calculations based on the perturbation approach, and many optimization iterations. Two phases of design with our approximation method used only 6.4 minutes and 1357 actual simulations. The final yield was 79.7%. In all cases the yield values were estimated from 1000 Monte Carlo samples using exact circuit simulations. For the purpose of clarity, the CPU time needed for this yield estimation is not included in the aforementioned CPU times. Interestingly, for this example, our method not only presented greatly reduced computational effort as compared with the actual simulation approach, but also reached a higher final yield. It suggests that accurate, time-consuming exact circuit simulation does not necessarily result in a better final yield.

While not illustrated in the table, we have used this example to compare the efficiency of our method w.r.t. the original maximally flat approach (Biernacki and

TABLE 3.1

COMPARISON OF STATISTICAL DESIGN OF A LOW-PASS FILTER WITH AND WITHOUT QUADRATIC APPROXIMATIONS

Component	Nominal Design	Exact Simulation		Quadratic Approximation	
		Phase 1	Phase 2	Phase 1	Phase 2
x_i	x^0	x^1	x^2	x^3	x^4
x_1	0.22510	0.22572	0.22512	0.22266	0.21669
x_2	0.24940	0.24903	0.24944	0.25045	0.25131
x_3	0.25230	0.25269	0.25276	0.25268	0.25083
x_4	0.24940	0.24908	0.24882	0.25028	0.24067
x_5	0.22510	0.22568	0.22594	0.22335	0.22120
x_6	0.21490	0.21589	0.21658	0.22163	0.23347
x_7	0.36360	0.36313	0.36275	0.36291	0.37008
x_8	0.37610	0.37625	0.37698	0.37938	0.37217
x_9	0.37610	0.37633	0.37561	0.37156	0.38529
x_{10}	0.36360	0.36313	0.36305	0.36226	0.37232
x_{11}	0.21490	0.21587	0.21674	0.22168	0.21893
Yield†	54.0%	61.7%	63.7%	70.2%	79.7%
Yield††	54.0%			74.0%	84.5%
Number of Outcomes Used for Optimization		200	200	200	200
Starting Point		x^0	x^1	x^0	x^3
Number of Simulations		48000	31200	529	828
Number of Iterations		9	7	10	19
CPU(VAX 8600)		96.3min.‡	62.4min.‡	2.5min.	3.9min.
CPU(MicroVAX)		481min.	312min.	12.3min.	19.5min.‡

CPU times do not include yield estimation based on actual simulation.

† The yield is estimated using exact simulation and 1000 outcomes.

†† The yield is estimated using quadratic approximation and 200 outcomes used in design.

‡ The CPU time is approximately given by assuming that the speed ratio of VAX 8600 to MicroVAX is 5.

Styblinski 1986). Using the same base points and employing the same scheme of rebuilding the models at each iteration, the original approach required approximately 5.3 and 8.3 minutes for the same two phases that our method took 2.5 and 3.9 minutes to finish. It should be noted that in both cases the CPU time needed to build and/or to evaluate the model constitutes only a fraction of the overall time, thus the remaining portions are common to the two approaches.

3.5 GRADIENT QUADRATIC APPROXIMATION SCHEME

In the previous presentation, we have applied quadratic approximation to circuit response functions. In this section, we utilize an efficient quadratic approximation scheme to replace the expensive repeated circuit simulations and gradient evaluations, in order to speed up the process. The novelty of this utilization is that not only circuit performance functions, but also their gradients are approximated. In a gradient-based optimization procedure, such as the one-sided ℓ_1 centering approach (Bandler and Chen 1989), gradient information is critical in determining the direction for optimization iterations to follow. Higher gradient accuracy will improve the overall performance of the optimization process.

3.5.1 Quadratic Approximations to Responses and Gradients

In most approximation approaches for statistical design, such as Biernacki and Styblinski (1986) and Biernacki, Bandler, Song and Zhang (1989), only circuit responses are modeled by quadratic functions. The gradients of the responses are either not used or their approximate values are calculated by differentiating the quadratic approximate responses. To further improve the performance of the

gradient-based yield-driven optimization, more accurate gradients are preferable.

Consider a response function with n variables. The gradient of the response is a vector of functions of the same n variables, each of the functions being the partial derivative of the response w.r.t. one *designable* variable. In yield optimization we typically deal with three types of variables, namely, n_{DS} designable variables x_{DS} with statistics, n_D designable variables x_D without statistics, and n_S non-designable variables x_S which are subject to statistical variations. Suppose that there are k responses, R_i, $i = 1, 2, \ldots, k$. The gradients of the responses with respect to the designable variables are

$$\nabla R_i = \left[\left[\frac{\partial R_i}{\partial x_{DS}^0} \right]^T \left[\frac{\partial R_i}{\partial x_D^0} \right]^T \right]^T, \tag{3.18}$$

where x^0 stands for the nominal values and the dimension of the gradient vector is $(n_{DS} + n_D)$.

For yield-driven design, circuit responses and their gradients have to be evaluated at a number of statistical outcomes. Each statistical outcome is generated in a $(n_{DS} + n_S)$-dimensional space according to a known statistical distribution and can be expressed as

$$\begin{bmatrix} x_{DS} \\ x_S \end{bmatrix} = \begin{bmatrix} x_{DS}^0 \\ x_S^0 \end{bmatrix} + \begin{bmatrix} \Delta x_{DS} \\ \Delta x_S \end{bmatrix}, \tag{3.19}$$

where Δx_{DS} and Δx_S are outcome specific deviates from the nominal values. Because of a large number of statistical outcomes needed for a meaningful yield estimate the main saving of the computational effort is achieved by building the models in the $(n_{DS} + n_S)$-dimensional space of the statistical variables (3.19). In other words, we

consider (3.19) as the vector x in (3.1), or

$$x = \begin{bmatrix} x_{DS} \\ x_S \end{bmatrix}. \tag{3.20}$$

Locality of statistical spreads assures a good level of model accuracy. The models are built for the current (optimization specific) nominal point and utilized for as many statistical outcomes as desired. In addition to the response functions, each entry to the gradient vectors can be approximated by a separate quadratic function in a similar manner as the response functions are. For k responses, thus, there are totally $k\times(1+n_{DS}+n_D)$ functions to be approximated, i.e.,

$$\begin{bmatrix} R_1 \\ \nabla R_1 \\ R_2 \\ \nabla R_2 \\ \cdot \\ \cdot \\ \cdot \\ R_k \\ \nabla R_k \end{bmatrix}. \tag{3.21}$$

It should be pointed out that, if the adjoint technique is used, the gradient can be available at a low additional cost to the circuit simulation, and can be returned from the simulator regardless of whether it is utilized or not. Therefore, the proposed method can not only utilize information that would otherwise be lost, but also allows for reduction of the model dimensionality by n_D, as is clearly seen from (3.21).

A general-purpose circuit design program, called McCAE, is under continuous development in the Simulation Optimization Systems Research Laboratory (SOSRL) of McMaster University. This program uses proprietary modules of Optimization

Systems Associates Inc. In this program, the gradient information has been made available. Implementation II of the generalized ℓ_1 centering approach is used. We apply our quadratic approximation technique to gradient functions as well. $2(n_{DS} + n_S) + 1$ base points, defined by (3.10) and (3.13), are used. An interface has been developed for the response and gradient approximation module which is very flexible in dealing with different types of variables involved in yield-driven design. Additional to response functions, their gradient functions are also taken as candidates to be approximated. The same set of $2\times(n_{DS} + n_S) + 1$ base points, defined by (3.10) and (3.13), is used both for response functions and for their gradient functions. Again, the reference point is defined as the nominal point. At each iteration, a set of quadratic models for both responses and gradients is built. The models are evaluated for all outcomes. The resulting quadratic model for the gradient is more accurate than the one that could be obtained by differentiating the quadratic model of the response, because the partial second-order information is incorporated in the model.

3.5.2 A 13-Element Low-Pass Filter

The low-pass filter (Wehrhahn and Spence 1984) shown in Fig. 3.2 is considered. The circuit must meet the specifications: insertion loss less than 0.4dB at the angular frequencies

{0.25, 0.27, 0.29, 0.31, 0.33, 0.67, 0.69, 0.71, 0.73, 0.75, 0.90,

0.905, 0.91, 0.92, 0.93, 0.978, 0.981, 0.984, 0.986, 0.988, 1},

and greater than 49dB at

{1.04569, 1.056, 1.059, 1.063, 1.067, 1.071, 1.115}.

There are 13 design variables. A normal distribution with 0.5% standard

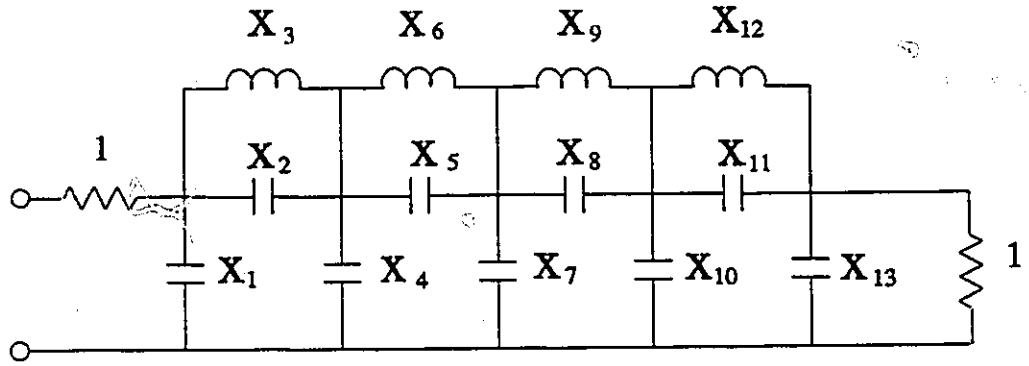

Fig. 3.2. Circuit schematic of the LC 13-element filter (Wehrhahn and Spence 1984).

deviation is assumed for all variables. The starting point is the optimal minimax solution, which has an estimated yield of 33.4%. To illustrate the efficiency of the new quadratic approximation approach, we solve the problem using both approximate simulations from the quadratic model and exact simulations. The final yields for both approaches are 75.6 and 80.7%. Computational details are given in Table 3.2. CPU times for the two designs were 7 and 30 minutes, respectively.

3.5.3 A Two-Stage GaAs MMIC Feedback Amplifier

We consider a two-stage 2-6GHz GaAs MMIC feedback amplifier (Vendelin, Pavio and Rohde 1990). The specifications are a small-signal gain of 8dB±1dB, VSWR at the input port of less than 2, and VSWR at the output port less than 2.2. The circuit and the equivalent circuit model for the FET are shown in Figs. 3.3 and 3.4. It is intended to manufacture high-volume, high-yield, and, consequently, low-cost microwave amplifiers. The size of the IC has a strong effect on the cost. Therefore, we consider the mean values of most capacitors as fixed to keep the size of the chip reasonable. The mean value of the gate width is fixed because of the assumed FET process, but a 3% standard deviation is allowed. We assume that the bias circuit is well designed such that changes of bias resistor values do not strongly influence the RF response around the operating point. Therefore, no tolerances are assigned to the resistors. Two feedback resistors and a forward capacitor are chosen as design variables. A standard deviation of 2% is assumed for the design variables. For nominal values and standard deviations of other elements, see Table 3.3. Correlations among elements are not considered in this circuit.

The first step in the entire optimization procedure is to find a minimax

TABLE 3.2

YIELD OPTIMIZATION OF THE LC 13-ELEMENT FILTER WITH AND WITHOUT QUADRATIC APPROXIMATIONS

Parameter	Initial	Solution†	Solution††
x_1	0.2088	0.2145	0.2205
x_2	0.03594	0.03642	0.03929
x_3	0.1822	0.1800	0.1775
x_4	0.2340	0.2347	0.2266
x_5	0.2424	0.2426	0.2556
x_6	0.08776	0.08702	0.08426
x_7	0.1333	0.1290	0.1234
x_8	0.3549	0.3535	0.3551
x_9	0.06477	0.06496	0.06481
x_{10}	0.1674	0.1625	0.1561
x_{11}	0.1422	0.1435	0.1498
x_{12}	0.1140	0.1120	0.1098
x_{13}	0.1433	0.1414	0.1303
Yield Estimate	33.4%	75.6%	80.7%
Number of Simulations		3780	18200
CPU*		7min.	30min.

† The solution after one phase of yield optimization with quadratic approximations.
†† The solution after one phase of yield optimization with exact simulations and numerical gradients.
* On the Sun SPARCstation 1.

Comments: Normal distribution of $\sigma = 0.5\%$ is assumed for all parameters. 100 outcomes are used in the optimization.
1000 outcomes are used in the yield estimation.

Parameters are scaled down by the factor 2π, e.g., the actual element value of x_1 is $2\pi \times 0.2088$.

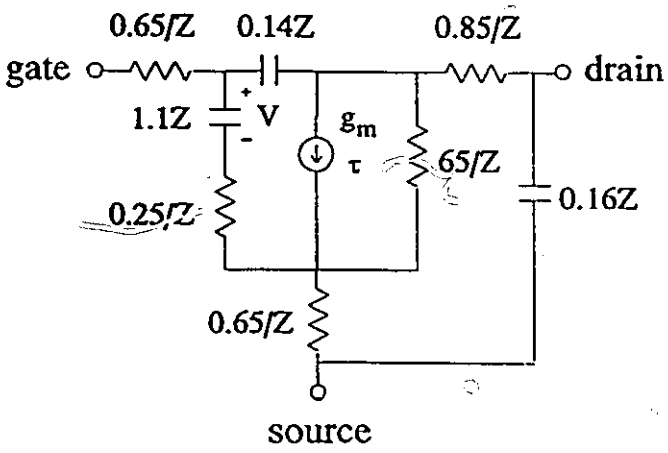

Fig. 3.3 Normalized GaAs MESFET model (Vendelin, Pavio and Rohde, 1990). Z is the gate width in millimeters. $g_m = 0.17Z$, and $\tau = 2.5$ps. All resistors are in ohms. All capacitors are in picofarads.

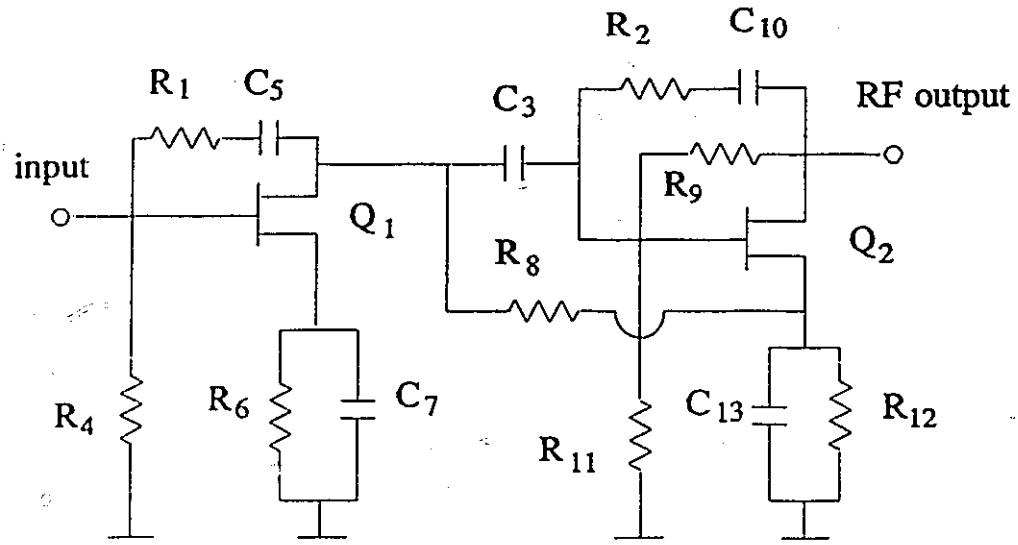

Fig. 3.4 A two-stage amplifier (Vendelin, Pavio and Rohde, 1990).

TABLE 3.3

PARAMETER VALUES AND TOLERANCES FOR THE MMIC AMPLIFIER

Element Parameter	Mean Value	Standard Deviation
$Z(\mu m)$	300	3%
$R_4(\Omega)$	400	0%
$C_5(pF)$	4	2%
$R_6(\Omega)$	20	2%
$C_7(pF)$	10	2%
$R_8(\Omega)$	145	2%
$R_9(\Omega)$	2200	0%
$C_{10}(pF)$	4	2%
$R_{11}(\Omega)$	6000	0%
$R_{12}(\Omega)$	500	2%
$C_{13}(pF)$	10	2%

TABLE 3.4

YIELD OPTIMIZATION OF THE MMIC AMPLIFIER WITH AND WITHOUT QUADRATIC APPROXIMATIONS

Parameter	Initial	Solution†	Solution††
R_1	201.02	207.63	207.73
R_2	504.82	627.94	630.53
C_3	5.3501	2.7742	2.7563
Yield Estimate	32.1%	77.8%	77.3%
Number of Simulations		1380	6400
CPU*		9min.	39min.

† The solution after one phase of yield optimization with quadratic approximations.
†† The solution after one phase of yield optimization with exact simulations and numerical gradients.
* On the Sun SPARCstation 1.

Comments: 100 outcomes are used in the optimization. 1000 outcomes are used in the yield estimation.

solution as the starting point for yield-driven design. The minimax solution is found and listed in Table 3.4. The yield estimate at this point is 32.1%. Two yield-driven optimization processes are carried out with and without the new quadratic approximation to both the responses and their gradients. Two solutions and the final yield are given in Table 3.4. The actual yields based on a Monte Carlo analysis of 1000 outcomes are 77.8% and 77.3%, respectively. The CPU times are 9 and 39 minutes.

3.6 CONCLUDING REMARKS

In this chapter we have presented a highly efficient quadratic approximation technique. The new approach takes advantage of the maximally flat interpolation and of a fixed pattern of base points, thus substantially reducing the computational effort and required storage. A set of extremely simple formulas to calculate model coefficients has been derived. The elegance of this approach is its conciseness and applicability. The very strong impact of our approach on the feasibility of statistical design of larger circuits should not be underestimated.

This approximation approach has been applied either to circuit response functions only, or to responses and their gradients simultaneously. For the first case, the results of a statistical design example has proven very efficient. For the second case, it is found that our approach can be especially suitable for gradient-based yield-driven design, making the model more accurate and robust. A standard test problem and an MMIC amplifier design illustrate the merits of this implementation.

It should also be noted that our approach is suitable for a large variety of applications where a large number of expensive simulations is involved.

4

SUPERCOMPUTER-AIDED STATISTICAL DESIGN OF A LARGE SCALE CIRCUIT — A 5-CHANNEL MICROWAVE MULTIPLEXER

4.1 INTRODUCTION

The advantage of modern technology has been creating increasingly complex circuits. The demand for reliable design which leads to shorter development time, requires more and more accurate, at the same time, more and more computationally involved models for circuit elements and devices. Statistical design of large scale circuits presents a great challenge.

Design of microwave multiplexers is a large scale problem. Contiguous-band multiplexers consisting of multi-cavity filters distributed along a waveguide manifold are used in satellite communications. The problem of optimal design and manufacture of such circuits has been of significant, practical interest (Atia 1974; Chen, Assal and Mahle 1976, Chen 1983 and 1985) for a number of years. There has been systematic research to provide very comprehensive simulation, sensitivity analysis, and optimization design tools for these circuits (Bandler, Kellermann and Madsen 1985 and 1987, Bandler, Chen, Daijavad and Kellermann 1984 and 1988, Bandler, Daijavad and Zhang 1985 and 1986, and Bandler and Zhang 1987).

Bandler, Biernacki, Chen, Renault, Song and Zhang (1988) developed a combined approach to large scale circuit statistical design. They attacked this difficult problem by: (1) the use of supercomputers, (2) efficient approximation to circuit responses, and (3) the use of fast, dedicated simulation techniques.

In the manufacture of electrical circuits, tuning can be an essential and effective part of the production process to improve circuit performance and to increase the final manufacturing yield. Through adjustment of circuit components to meet design specifications, tuning counteracts collective effects caused by the tolerances inherent in the production process, model inaccuracy and other factors ignored in the design process. In general, tuning requires simultaneous adjustment of several elements to reach a satisfactory result. Even when the number of tunable elements is moderate, the effort made to find the right combination of elements and amounts to be adjusted is enormous. This leads to the most apparent problem with the tuning process: the enormous amount of time required. In the tuning process, the amount of tuning is also subject to imprecisions of adjustments and uncertainties of circuit models. This chapter also addresses the issue of the application of statistical design to tunable circuits with tolerances associated with both the fabricating process prior to tuning and tuning process itself.

This chapter is organized in the following way. We first provide a physical meaning of applying yield-driven design principles to design of small production volumes. Then special consideration is given to the application of yield-driven design to tunable circuits. We discuss the use of the supercomputer to deal with large scale circuits. Finally, a 5 channel multiplexer with 75 design and toleranced variables, 124 constraints and up to 200 statistically perturbed outcomes, is used to demonstrate the feasibility of yield optimization of large scale problems using our combined strategy.

4.2 YIELD-DRIVEN DESIGN FOR NON-MASSIVE PRODUCTION

Conventional yield-driven design is intended to increase yield of mass production of electronic circuits. In some cases, however, a relatively small number of circuits are requested by customers. Then the term yield itself is no longer mathematically (more precisely, statistically) justifiable because of the small number of outcomes. Under such circumstances, the major concern of the design is not the yield, but the probability of producing satisfactory circuits at the lowest cost. Tolerances, inevitably inherent in the manufacturing process, are still a critical factor affecting the final outcomes. The logical result of applying the yield-driven design methodology to circuit production on a small scale is the improved probability of obtaining circuits meeting specifications.

The formal mathematical justification is based on the original definition of yield

$$Y(\phi^0) = \text{Probability of } \{ \phi^i \in R_a \}, \tag{4.1}$$

where R_a is the acceptable region and fixed tolerances given by ϵ are assumed. Then, the utilization of ordinary yield-driven design techniques can increase the probability of manufactured circuit outcomes to meet design specifications.

4.3 YIELD-DRIVEN DESIGN OF TUNABLE CIRCUITS

4.3.1 Tuning and Its Tolerances

To apply the available yield-driven design approach to tunable circuits, a number of aspects should be identified. In the following, tolerances inherent in the circuit fabrication process before tuning are referred to as fabrication tolerances, and tolerances in the tuning process as to tuning tolerances. Here, we focus our attention

on the case of fixed fabrication tolerances, tuning region, and assumed tuning tolerances. We assume that the fabrication tolerances are symmetrical and that tuning is two-way and symmetrical. Let

$$\epsilon = \left[\epsilon_1 \;\; \epsilon_2 \;\; \cdots \;\; \epsilon_N \right]^T, \tag{4.2}$$

$$t = \left[t_1 \;\; t_2 \;\; \cdots \;\; t_N \right]^T, \tag{4.3}$$

and

$$\epsilon_t = \left[\epsilon_{t_1} \;\; \epsilon_{t_2} \;\; \cdots \;\; \epsilon_{t_N} \right]^T \tag{4.4}$$

denote the fabrication tolerance extremes, maximal tuning amounts, and tuning tolerance extremes, respectively. For those elements that cannot be tuned, the corresponding tuning amounts in (4.3) are zero. Then, the associated tuning tolerances in (4.4) are also zero. We will assume that the parameters can be varied continuously and independently. ϕ^0 is used to denote the nominal design. The kth tunable outcome ϕ^k may be described by

$$\begin{aligned} \phi^k &= \phi^0 + \Delta\phi^k, \\ \Delta\phi^k &= \Delta\phi_\epsilon^k + t^k + \Delta t^k, \end{aligned} \tag{4.5}$$

where $\Delta\phi^k$ needs the following explanation: $\Delta\phi_\epsilon^k$ represents deviations due to the actual model uncertainties and fabrication tolerances, generated from

$$\Delta\phi_\epsilon^k = E\mu, \tag{4.6}$$

where

$$E \triangleq \begin{bmatrix} \epsilon_1 & & & & \\ & \epsilon_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \epsilon_N \end{bmatrix}, \qquad \mu \triangleq \begin{bmatrix} \mu_1 \\ \mu_2 \\ \cdot \\ \cdot \\ \cdot \\ \mu_N \end{bmatrix}, \quad -1 \le \mu_i \le 1, \quad i = 1, 2, \ldots, N;$$

t^k denotes possible postproduction tuning adjustments, taking form of

$$t^k = T\rho \,, \tag{4.7}$$

where

$$T \triangleq \begin{bmatrix} t_1 & & & & \\ & t_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & t_N \end{bmatrix}, \qquad \rho \triangleq \begin{bmatrix} \rho_1 \\ \rho_2 \\ \cdot \\ \cdot \\ \cdot \\ \rho_N \end{bmatrix}, \quad -1 \le \rho_i \le 1, \quad i = 1, 2, \ldots, N;$$

and Δt^k represents actual tuning tolerances, defined by

$$\Delta t^k = E_t \sigma \,, \tag{4.8}$$

where

$$E_t \triangleq \begin{bmatrix} \epsilon_{t_1} & & & & \\ & \epsilon_{t_2} & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \epsilon_{t_N} \end{bmatrix}, \qquad \sigma \triangleq \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \cdot \\ \cdot \\ \cdot \\ \sigma_N \end{bmatrix}, \quad -1 \le \sigma_i \le 1, \quad i = 1, 2, \ldots, N.$$

In these formulas, μ_i and σ_i are random numbers, generated according to their own statistics, ρ_i is the relative tuning amount. For ideally accurate tuning ($\Delta t^k = 0$) we can obtain $\Delta \phi^k = 0$ provided that the tuning ranges are large enough to accommodate

the spreads due to the fabrication tolerances. Then, for realistic (imprecise) tuning and sufficient tuning region, we obtain $\Delta\phi^k = \Delta t^k$. Outcomes of nontunable circuits can be considered as a special case of (4.5), where $\Delta\phi^k = \Delta\phi_\epsilon^k$.

If the yield optimization only takes manufacturing tolerances into account, we may face a very poor yield figure because many fabricated circuits cannot be tuned into the acceptable region due to existing tolerances. Consider a two-dimensional case shown in Fig. 4.1, where ϵ_1 and ϵ_2 are fabrication tolerances. If only the fabrication tolerances are considered in yield optimization, design tends to move the fabrication tolerance region, R_ϵ, to overlaps the acceptable region, R_a, as much as possible. Then we may have a solution ϕ^0. At this point, circuit outcomes fabricated can spread over the area R_ϵ. Assume that tuning on the second element is available and that the maximal tuning amount is t_2. Then, tuning on the second parameter tries to bring outcomes into the area $R_{\epsilon,\, t}(\phi^0, \bar{\epsilon})$, defined by the nominal point and

$$\bar{\epsilon} \triangleq \begin{bmatrix} \bar{\epsilon}_1 \\ \bar{\epsilon}_2 \\ \cdot \\ \cdot \\ \cdot \\ \bar{\epsilon}_N \end{bmatrix}, \tag{4.9}$$

where

$$\bar{\epsilon}_i \triangleq \max(0, \epsilon_i - t_i) \ . \tag{4.10}$$

Because of tuning tolerances, final outcomes (after tuning) do not all fall into $R_{\epsilon,\, t}$. Instead, they spread over the region

$$R_{\epsilon,\, t,\, t_\epsilon}(\phi^0, \overline{\epsilon}) \ , \tag{4.11}$$

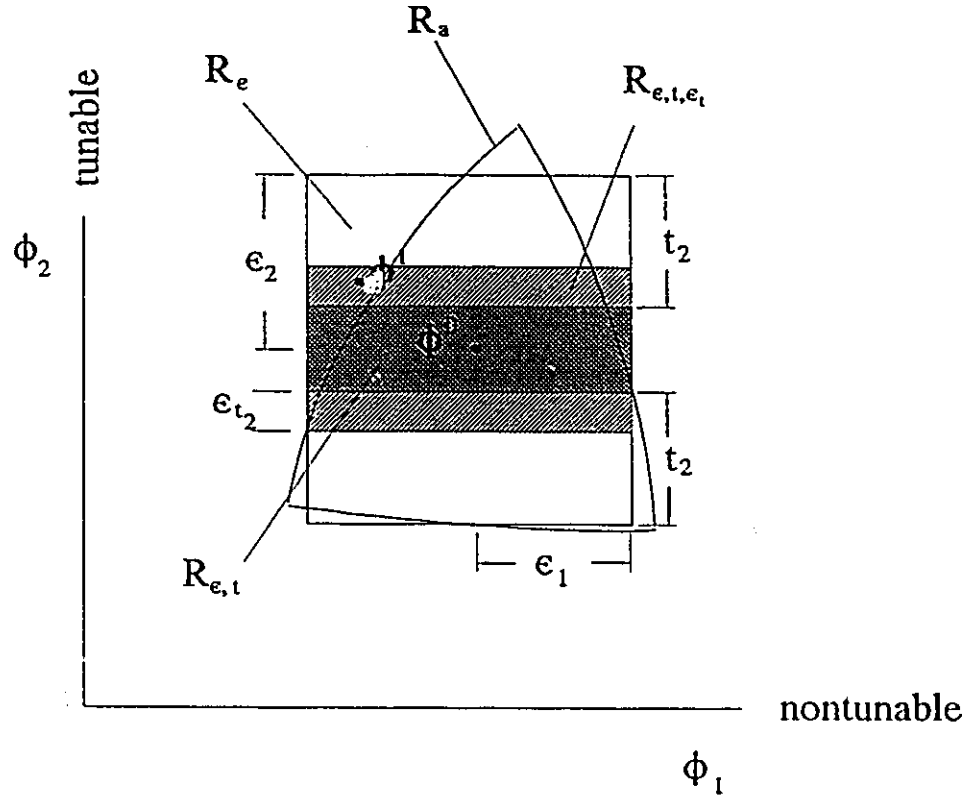

Fig. 4.1 Illustration of yield optimization with the fabrication tolerance region R_ϵ only. R_a is the acceptable region, $R_{\epsilon,t}$ the spread region with exact tuning, and $R_{\epsilon,t,t_\epsilon}$ is the actual spread region after toleranced tuning. Yield after tuning is $(R_a \cap R_{\epsilon,t,t_\epsilon})/R_{\epsilon,t,t_\epsilon}$.

where

$$\overline{\epsilon} = \tilde{\epsilon} + \epsilon_t \ .$$

Some of circuits may fall outside of R_a due to the tuning tolerances, indicating that those outcomes violate the design specifications. For instance, ϕ^1 shown in Fig. 4.1 is such a circuit.

4.3.2 Yield-Driven Design with Tuning Tolerances

It is natural to extend yield-driven design techniques to tunable circuits. Consider

$$\begin{aligned} \phi^k &= \phi^0 + \Delta\phi^k \\ \Delta\phi^k &= \overline{E}\theta \ , \end{aligned} \tag{4.12}$$

where

$$\overline{E} \triangleq \begin{bmatrix} \overline{\epsilon}_1 & & & \\ & \overline{\epsilon}_2 & & \\ & & \ddots & \\ & & & \overline{\epsilon}_N \end{bmatrix}, \quad \theta \triangleq \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_N \end{bmatrix}, \quad -1 \le \theta_i \le 1, \quad i = 1, 2, \ldots, N \tag{4.13}$$

If the combined outcome spread region, $R_{\epsilon, t, t_\epsilon}(\phi^0, \overline{\epsilon})$, is treated in the same way as the normal tolerance region, then yield-driven design will give a solution such that tunability is fully utilized to cancel any deviation caused by fabrication tolerances and, meanwhile, tuning tolerances are also accommodated. Fig. 4.2 presents such a formulation where tuning and tuning tolerances are considered. Unlike the formulation of Fig. 4.1, this procedure tends to move the nominal point such that the combined region $R_{\epsilon, t, t_\epsilon}(\phi^0, \overline{\epsilon})$ overlaps the acceptable region R_a as much as possible,

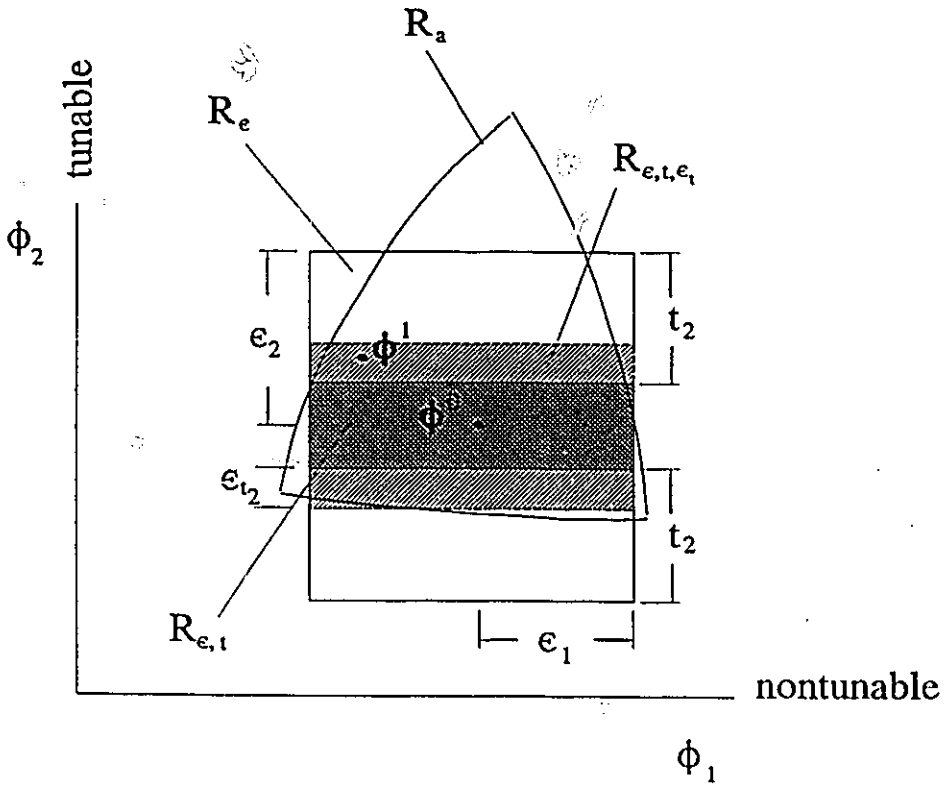

Fig. 4.2 Illustration of yield optimization with the combined spread region $R_{\epsilon,t,t_\epsilon}$. Yield after tuning is higher than that in Fig. 4.1.

giving a higher yield after tuning with the same tunable region and tuning tolerances. Now, ϕ^1 is well in the acceptable region.

The design of tunable circuits using (4.12) has two goals. One is to drive up the probability of obtaining circuits that exhibit good initial responses for the tuning process. The other is to increase the possibility of circuit outcomes satisfying specifications after tuning easy-to-tune elements. For the latter, we assign larger tuning tolerances to more difficult-to-tune elements in an effort to avoid painstaking, and tedious tuning. This should have an impact on the overall effort of tuning.

4.4 UTILIZING THE VECTOR PIPELINE SUPERCOMPUTER

4.4.1 Basics about the Vector Processor

The largest available computers, namely supercomputers, have impressive number-crunching and storage capabilities. One typical type of supercomputer is the vector pipeline computer.

One of the most fundamental mechanisms that makes a vector pipeline supercomputer very powerful is its machine architectures, which include vector instruction and pipelining. From the user's point of view, the program should be reorganized such that it best fits the machine architecture. Here, we concentrate only on this type of reorganization called program vectorization.

Program vectorization is the most important action to be taken to fully take advantage of a vector pipeline supercomputer. The vectorization is carried out by the compiler. What the compiler actually vectorize is do-loops. Specifically, it vectorizes innermost do-loops, so in any section of code with do-loops nested more than one level deep, the compiler will attempt to vectorize only the most deeply-nested loop.

In circuit simulation and optimization programs, a substantial amount of work is done by many vector and matrix manipulations, resulting in many possibly vectorizable do-loops. Successful vectorization of these programs will get dramatic performance improvement.

The following steps should be taken to make a program better vectorized. Often, we encounter a pair of do-loops nested for which the outer loop has a much greater range than the inner loop. We can rearrange the loops by exchanging them to gain efficiency. Keep the inner loop as simple as possible. Many extrinsic function calls, I/O statements, GOTO and IF statements, and index cross references inhibit vectorization. Certain preprocessing may be helpful to reorganize the code structure to achieve better vectorization results at the cost of more memory which usually is not a big problem with supercomputers.

4.2.2 Supercomputer-Aided Yield Optimization

Vector pipeline supercomputers have been used by many researchers in the circuit CAD field (Calahan 1979 and 1980, Vladimirescu and Pederson 1982, Yamamoto and Takahashi 1985). Specially, effort has been made to explore supercomputers for microwave circuit CAD (Rizzoli, Ferlito and Neri 1986, Rizzoli, Cecchetti and Lipparini 1986, Rizzoli and Neri 1988, Bandler, Biernacki, Chen, Renault, Song and Zhang 1988, Rizzoli et al 1991).

From the CAD viewpoint, yield-driven design problems concerning realistic microwave circuits are of very large scale. With existing design methods, the most powerful available computational tools, namely, supercomputers, can provide us with an effective means to deal with practical design problems.

4.4.3 The Cray X-MP/22 Environment

Yield-driven design has been carried out on the CRAY X-MP/22 in the Ontario Center for Large Scale Computation, located at the University of Toronto. Included in the mainframe are two central processing units (CPUs), each controlled by a 9.5 nanosecond (ns) clock, and 2 Million 64-bit words (16MB) of shared central memory. The Cray Operating System (COS) efficiently allocates system resources and controls the execution of user jobs on one or both of the CPUs. At the time the system ran under Version 1.15 of (COS). The FORTRAN compiler used is Cray's FORTRAN-77 (CFT77). This compiler performs optimizations such as automatic vectorization and instruction resequencing in order to most effectively utilize the X-MP hardware.

4.5. YIELD OPTIMIZATION OF A 5-CHANNEL MULTIPLEXER

This problem is a 5-channel 12GHz contiguous band microwave multiplexer consisting of multi-cavity filters distributed along a waveguide manifold (Bandler, Chen, Daijavad and Kellermann 1984). Fig. 4.3 illustrates the equivalent circuit of the multiplexer. Tuning is essential and expensive for multiplexers to satisfy the ultimate specifications. The goal of this design is to provide such a well-centered nominal design that the tuning process can be greatly easied.

General multiplexer optimal nominal design procedures using powerful gradient-based minimax and ℓ_1 algorithms have been described by Bandler, Kellermann and Madsen (1985 and 1987). The circuit simulation and sensitivity analysis aspect of the problem together with a number of examples of multiplexer optimization have been presented by Bandler, Chen, Daijavad and Kellermann (1984).

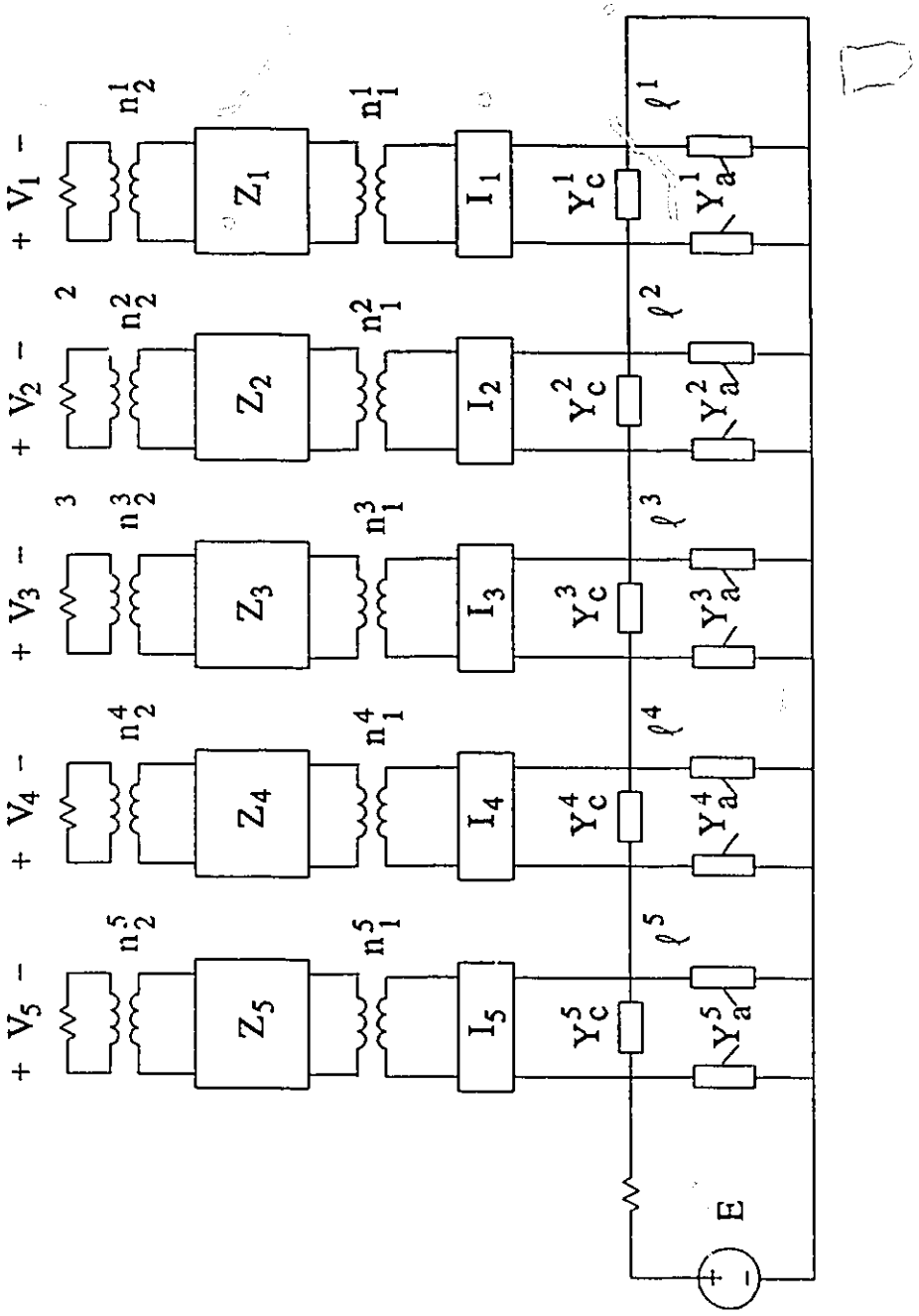

Fig. 4.3 Equivalent circuit of a 5-channel contiguous band multiplexer.

A novel approach to the exact simulation and sensitivity analysis of multiplexing networks has been derived by Bandler, Daijavad and Zhang (1986), based on the general branched cascaded network structure (Bandler, Daijavad and Zhang 1985). An automatic decomposition approach to optimization of large scale problems was applied to a 16-channel multiplexer design (Bandler and Zhang 1987). Recently, a 5-channel multiplexer has also been used to demonstrate a gradient approximation scheme used in gradient-based optimization (Bandler, Chen, Daijavad and Madsen 1988). A multiplexer design system has been developed in the Simulation Optimization Systems Research Laboratory, McMaster University.

Our yield optimization uses the generalized ℓ_1 centering algorithm described in Chapter 2. Even with the power of the supercomputer, such design using exact simulation would take a very long time to complete. The cost would be extremely high. To speed up the design process, we use the quadratic approximation scheme, presented in Chapter 3, to model the multiplexer responses. Therefore, only the simulation portion of the multiplexer design system is to be used. The gradient information will be provided using the quadratic models of the response functions.

The program is implemented on the CRAY X-MP/22. The multiplexer design system is tailored to obtain a dedicated simulator to perform fast circuit simulation. Then this part is connected to the statistical design driver, with an interface of quadratic approximation.

To fully explore the vector pipeline supercomputer, the program, including simulation, optimization and statistical outcome generation, is carefully reorganized by restructuring the program, reordering some DO loops, and redirecting input and output flows.

4.5.1 Design Variables and Specifications

As shown in Fig. 4. 3, each branch unit consists of an impedance inverter I_i, an input transformer with the transform ratio n_1^i, a multi-coupled cavity filter Z_i, an output transformer with the transform ratio n_2^i, and an output load with voltage V_i. The main cascade is a waveguide manifold with spacing l_i between two branches. Branches are connected to the main cascade through series junctions. The series junctions are assumed non-ideal. Each of the multi-coupled cavity filters is determined by 12 independent coupling coefficients.

In order to take the appropriate tolerances into account, specifications were chosen to be 10dB for the common port return loss and for the individual channel stopband insertion losses, resulting in 124 nonlinear constraint functions. Design variables included 60 couplings, 10 input and output transformer ratios and 5 waveguide spacings. Tolerances of 5% were assumed for the spacings, and tolerances of 0.5% for the remaining variables. The starting point was the solution of the conventional minimax nominal design w.r.t. specifications of 20dB. The corresponding responses are shown in Fig. 4.4. The estimated yield w.r.t. the specifications of 10dB at this point was 75%.

Careful and time-consuming tuning is essential for multiplexers. Any adjustment involving physical disconnection of the structure is particulary expensive. For the sake of illustration we focus attention in this example on the waveguide spacings as being expensive adjustments to make. Therefore, the larger tuning tolerances are assumed for the spacings. We consider total tolerances of 5% for the spacings, and 0.5% for the couplings and the transformer ratios.

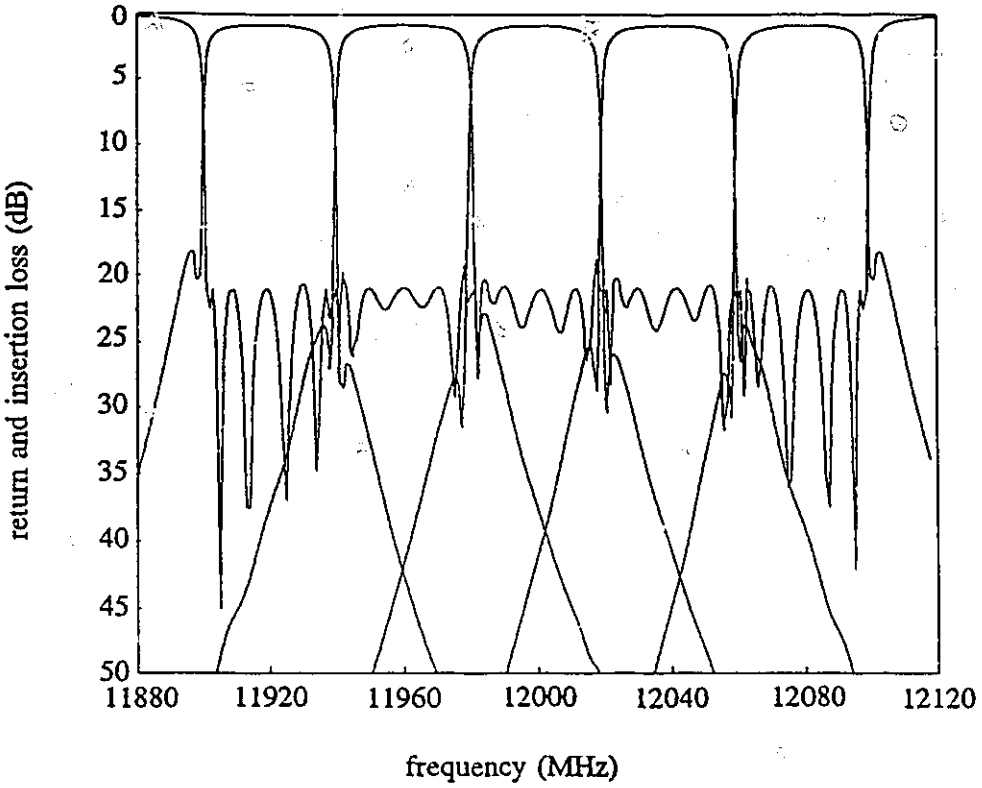

Fig. 4.4 Return and insertion loss of the 5-channel multiplexer at the minimax solution.

4.5.2 Design Procedure and Results

The process consists of 4 phases as shown in Table 4.1. At the beginning of each phase, a set of quadratic models corresponding to 124 responses is constructed. These models, then, are used for all outcomes in the entire phase. Notice that this is different from the cases in the last chapter where quadratic models are rebuilt for each optimization iteration. To build quadratic models, 151 simulations are performed at the base points for each phase, leading to a total 604 exact simulations for the entire design process. More and more outcomes, from 50 to 200, are used in four consecutive design phases. A total of 20 optimization iterations are involved. Without quadratic modeling, 2500 exact simulations and sensitivity analyses would have been required (if the same number of optimization iterations are necessary).

Four phases took totally 69.5 seconds on the CRAY X-MP/22 to reach a 90% estimated yield. The estimated yield is gradually increased in the first three phases and becomes almost steady in the last phase. To study the collective performance of the circuit, the return losses of all 3000 outcomes and of the satisfactory outcomes are contained within the envelopes in Figs. 4.5 and 4.6, respectively. A very small portion above the specification in Fig. 4.5 reflects the probability of near 10% to produce a bad circuit.

This approach allows us to handle this large optimization problem (with 75 toleranced variables, 124 constraints and up to 200 statistically perturbed circuits) in acceptable CPU time.

TABLE 4.1

STATISTICAL DESIGN OF A 5-CHANNEL MULTIPLEXER USING QUADRATIC APPROXIMATION

	Phase 1	Phase 2	Phase 3	Phase 4
Starting Point of the Phase	Nominal Design	Solution of Phase 1	Solution of Phase 2	Solution of Phase 3
Initial Yield*	75.0%	81.0%	84.3%	90.0%
Initial Yield**	56.3%	69.0%	69.3%	92.0%
Number of Outcomes Used for Optimization	50	100	150	200
Number of Iterations	4	6	6	4
Final Yield*	81.0%	84.3%	90.0%	90.3%
Final Yield**	77.3%	77.3%	91.3%	94.0%
CPU Time (CRAY X-MP/22)	16.5s	17.6s	17.8s	17.6s

* Yield estimated using actual simulation.
** Yield estimated using quadratic approximation.

CPU times do not include yield estimations based on actual simulation.
All yields are estimated using 300 samples.

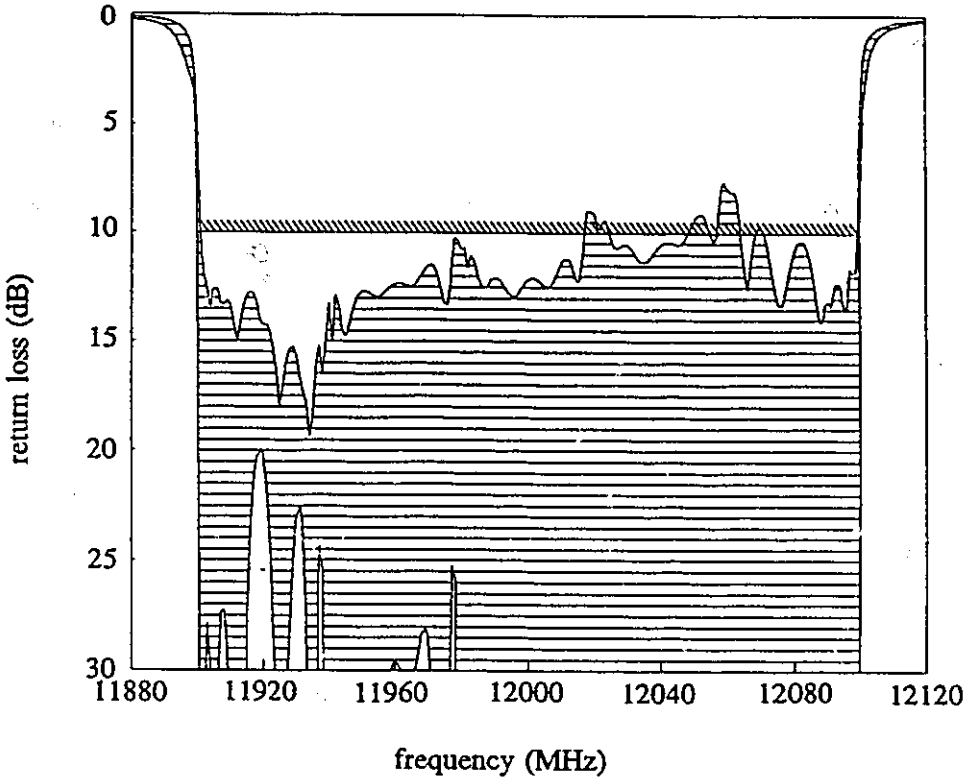

Fig. 4.5 Return loss envelope of 3000 5-channel multiplexer circuit outcomes after yield optimization.

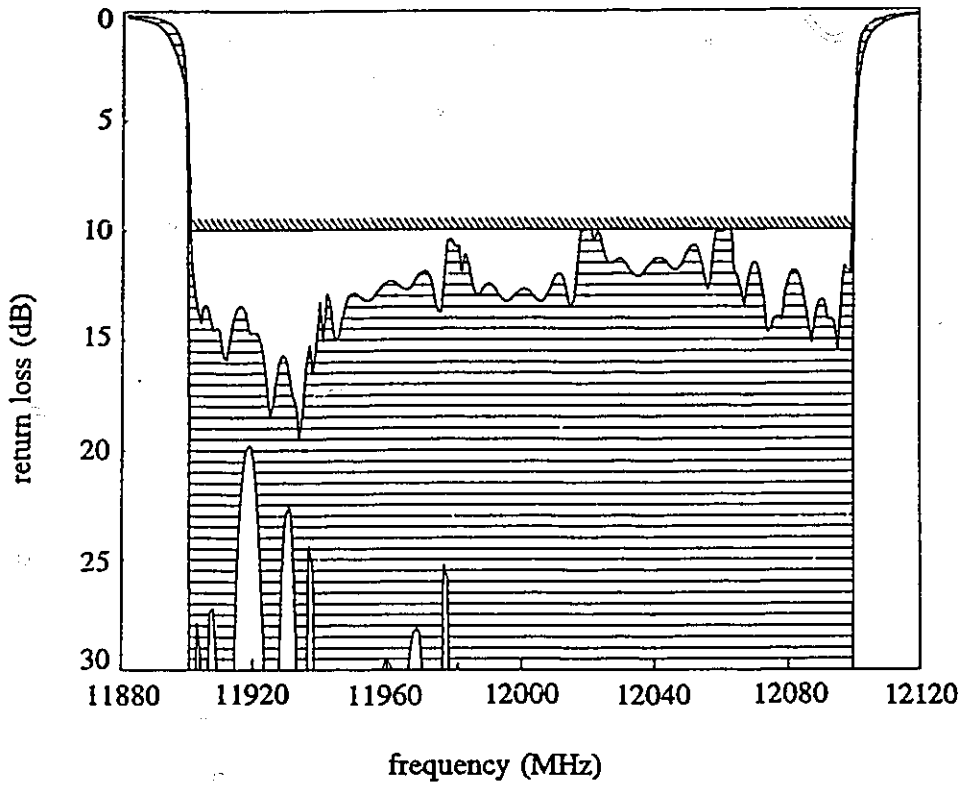

Fig. 4.6 Return loss envelope of satisfactory 5-channel multiplexers among 3000 circuit outcomes after yield optimization.

4.6 CONCLUDING REMARKS

In this chapter, we have shown that yield-driven design techniques can well be applied to non-massive production and tunable circuits. We have identified a number of aspects related to statistical design of tunable circuits with tuning tolerances. Some discussion addresses the use of one type of supercomputer, namely, the vector pipeline supercomputer. To cope with statistical design of large scale circuits, supercomputers and very efficient approximation techniques have been used to cooperate with a powerful yield-driven approach. To demonstrate the feasibility of such yield optimization of large scale microwave circuits, we have successfully accomplished the yield optimization of a 5 channel multiplexer with 75 design and toleranced variables, 124 constraints and up to 200 statistically perturbed circuits. No microwave circuit design optimization of this type and on this scale has ever been reported.

5

NONLINEAR CIRCUIT YIELD OPTIMIZATION WITH GRADIENT APPROXIMATIONS

5.1 INTRODUCTION

Many engineering applications require the use of nonlinear circuits, such as switches, oscillators, and mixers. The rapid development of monolithic microwave integrated circuits (MMICs) has made possible the integration of these nonlinear circuits within one die. This demands efficient and effective CAD tools to design nonlinear circuits. Not only must the tools produce valid nominal designs, but they must ensure design manufacturability and satisfactory yield. For MMIC technology, in particular, yield-driven, cost-effective design is vital to commercial competitiveness.

Rizzoli, Lipparini and Marazzi (1983) attempted to accomplishing nominal design of nonlinear circuits. Their approach combines the harmonic balance (HB) simulation and design into one optimization loop by considering design variables and HB state variable as optimization variables simultaneously. The circuit does not have to be completely solved by HB during optimization until the final solution is reached. However, this approach is not compatible with most yield-driven approaches since yield-driven design requires actual circuit responses to calculate yield as the objective. To meet this new and very difficult challenge, a number of important points must be incorporated in yield-driven design of nonlinear circuits:

(1) An efficient HB simulation scheme.

(2) High-speed gradient calculation if gradient-based optimization is used.

(3) Well-modeled nonlinear passive and active devices, such as FETs, diodes, and transmission lines, and valid statistical models for circuit components, especially for active devices.

(4) A well-structured over all design process consolidating the above points.

This chapter is organized as follows. We start by reviewing an efficient simulation method suitable for microwave nonlinear circuits, namely, the HB method (HB). Instead of working on very abstract and intricate formulations for the general HB approach, we use a simple circuit, a one-FET circuit, to illustrate how the HB method works. The HB method is implemented with exact Jacobian matrices for fast convergence and improved robustness. Next we introduce specifications and errors for nonlinear circuit yield optimization, and formulate the yield-driven design problem for nonlinear circuits. We offer an approach to efficient yield-driven optimization of nonlinear microwave circuits with statistically characterized devices (Bandler, Zhang, Song and Biernacki 1989). This approach utilizes a powerful and robust one-sided ℓ_1 optimization algorithm for design centering (Bandler and Chen 1988). The effective gradient approximation technique (Bandler, Chen, Daijavad and Madsen 1988) is adopted. An extension of this approach for statistical design is provided. Independent and/or correlated normal distributions and uniform distributions describing large-signal FET model parameters and passive elements are fully accommodated.

The yield optimization of a microwave frequency doubler with a large-signal statistically simulated FET model is successfully carried out. The performance yield is increased from 40% to 70%. We believe that this is the first demonstration of yield

optimization of nonlinear circuits operating under large-signal steady-state periodic or almost periodic conditions.

5.2 THE HARMONIC BALANCE SIMULATION TECHNIQUE

A quick review of the recently developed HB method is worthwhile because it is a key part of our yield-driven nonlinear circuit design approach. The formulation will also help us derive the sensitivity analysis in the next chapter. The HB method is a hybrid time- and frequency-domain approach which allows all the advantages of time-domain nonlinear device models, such as FETs and diodes, and the power of steady-state frequency-domain techniques for lumped and distributed circuit elements, such as microstrip lines and striplines. A variety of attempts has been made to improve the efficiency and versatility of HB by Kundert and Sangivanni-Vincentelli (1986), Rizzoli, Lipparini and Marazzi (1983), Curtice (1987), and Gilmore (1986), and many others.

The key concept in the HB method is to decompose the circuit into two parts, linear and nonlinear subnetworks. Therefore, time- and frequency-domain analyses can be performed on the nonlinear and linear subnetworks, respectively. To illustrate how the HB method can be used to simulate a nonlinear circuit, we consider a typical one-FET circuit shown in Fig. 5.1. The circuit is decomposed into the linear part, N_L, and the nonlinear part, N_{NL}, as shown in Fig. 5.2. We assume that the circuit is excited by a periodic stimulus with fundamental frequency ω.

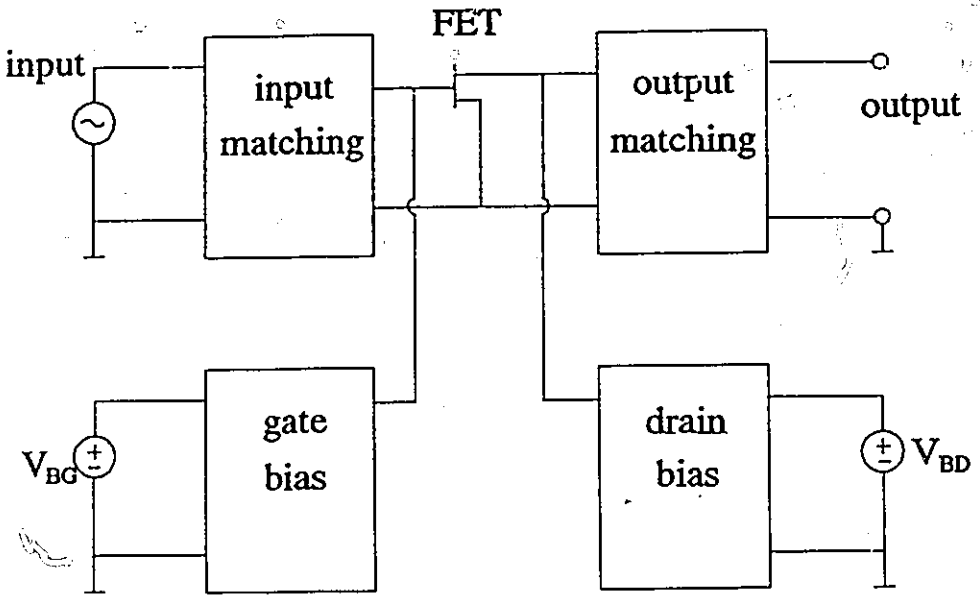

Fig. 5.1 Schematic representation of the one-FET circuit.

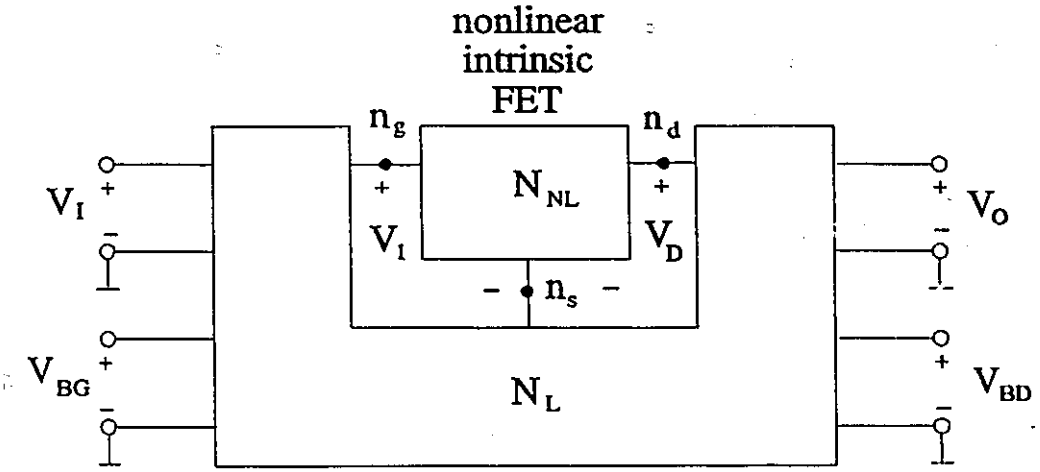

Fig. 5.2 Decomposition of the circuit in Fig. 5.1. N_L and N_{NL} stand for the linear and nonlinear subnetworks.

5.2.1 The Nonlinear Model and Its Time-Domain Simulation

The FET used is a modified Materka and Kacprzak model (Materka and Kacprzak 1985, Super-Compact 1987) shown in Fig. 5.3. The nonlinearities inherent in the intrinsic part are described by

$$i_D = F[v_G(t - \tau),\ v_D(t)](1 + S_s\frac{v_D}{I_{DSS}}),$$

$$F(v_G,\ v_D) = I_{DSS}\left[1 - \frac{v_G}{v_{p0} + \gamma v_D}\right]^{(E + K_E v_G)} \cdot \tanh\left[\frac{S_1 v_D}{I_{DSS}(1 - K_g v_G)}\right],$$

$$i_G = I_{G0}[\exp(\alpha_G v_G) - 1],$$

$$i_B = I_{B0}\exp[\alpha_B(v_D - v_1 - V_{BC})],$$

$$\begin{cases} R_1 = R_{10}(1 - K_R v_G), \\ R_1 = 0, \quad \text{if } K_R v_G \geq 1, \end{cases}$$

$$\begin{cases} C_1 = C_{10}(1 - K_1 v_G)^{-1/2}, \\ C_1 = C_{10}\sqrt{5} + C_{1s}, \quad \text{if } K_1 v_G \geq 0.8, \end{cases}$$

and

$$\begin{cases} C_F = C_{F0}[1 - K_F(v_1 - v_D)]^{(-1/2)}, \\ C_F = C_{F0}\sqrt{5}, \quad \text{if } K_F(v_1 - v_D) \geq 0.8, \end{cases}$$

where v_1, v_G, and v_D are controlling voltages as indicated in Fig. 5.3, the remaining coefficients being model parameters. Notice that such decomposition allows two of the controlling voltages, v_1 and v_D, to appear at the interfacing port.

Suppose that the controlling voltages can be expressed as

$$v_k(t) = \sum_{h=0}^{H} V_k(h)e^{jh\omega t}, \qquad k = 1,\ G,\ D, \tag{5.1}$$

where H is the highest harmonic considered and $V_k(h)$ are complex coefficients of the

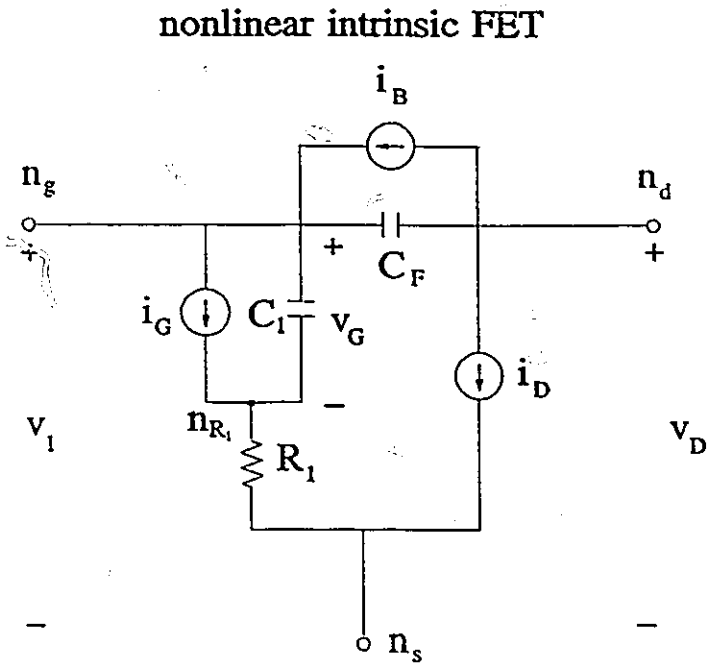

Fig. 5.3 The nonlinear intrinsic FET model.

Fourier series. Then we define

$$V_k = \begin{bmatrix} V_k(0) \\ V_k(1) \\ \cdot \\ \cdot \\ \cdot \\ V_k(H) \end{bmatrix}, \quad k = 1, G, D. \tag{5.2}$$

The currents from controlled sources, the current through the nonlinear resistor, and the charges on the capacitors are functions of V_1, V_G and V_D, that is,

$$i_k(V_G, V_D), \quad k = D, G, B$$

$$i_{R_1}(V_1, V_G) = \frac{[v_1(V_1) - v_G(V_G)]}{R_1(V_G)}$$

$$q_k(V_1, V_G, V_D), \quad k = 1, F$$

The calculation of the charges on the two nonlinear capacitors involves integration. The frequency-domain expressions of these currents and charges can be obtained by the Fourier transformation. They are of the forms

$$\begin{aligned} I_k(V) &= \mathcal{F}[i_k(V)], \quad k = D, G, B, R_1, \\ Q_k(V) &= \mathcal{F}[q_k(V)], \quad k = 1, F, \end{aligned} \tag{5.3}$$

where $\mathcal{F}$ stands for Fourier transform, I_k and Q_k are vectors consisting of the complex coefficients of the corresponding Fourier serieses, and V is defined by the following

$$V = \begin{bmatrix} V_1 \\ V_G \\ V_D \end{bmatrix}. \tag{5.4}$$

5.2.2 Linear Subcircuit Simulation

Bandler, Ye and Song (1990) proposed an efficient approach to construct the multiport matrix to accelerate HB analysis. Here we use this approach to deal with the circuit in Fig. 5.1.

Consider the linear subnetwork in Fig. 5.2. There are five ports, a port connecting with the intrinsic FET, an output voltage port, an input excitation port, and two DC bias ports. At a given frequency, the nodal equation can be written as

$$Y \begin{bmatrix} V_{n_g} \\ V_{n_d} \\ V_{n_s} \\ V_I \\ V_O \\ V_{BG} \\ V_{BD} \\ V_{n_1} \\ \cdot \\ \cdot \\ V_{n_b} \end{bmatrix} = \begin{bmatrix} I_{n_g} \\ I_{n_d} \\ I_{n_s} \\ I_I \\ I_O \\ I_{BG} \\ I_{BD} \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \tag{5.5}$$

where V_{n_k}, $k = g, d, s$, are node voltages at the three terminals of the interfacing port, and V_{n_i}, $i = 1, 2, \ldots, b$, are internal node voltages in the linear subnetwork.

First, we express Y in the form

$$Y = \left[y_{n_g} \; y_{n_d} \; y_{n_s} \; y_I \; y_O \; y_{BG} \; y_{BD} \; y_{n_1} \cdots y_{n_b} \right], \tag{5.6}$$

where each entry is a column vector. By adding and, at the same time, subtracting $y_{n_g} V_{n_s}$ and $y_{n_d} V_{n_s}$ to and from the right hand side of (5.5), we can rewrite it as

$$Y' \begin{bmatrix} V_{n_g} - V_{n_s} \\ V_{n_d} - V_{n_s} \\ V_{n_s} \\ V_I \\ V_O \\ V_{BG} \\ V_{BD} \\ V_{n_1} \\ \cdot \\ \cdot \\ V_{n_b} \end{bmatrix} = \begin{bmatrix} I_{n_g} \\ I_{n_d} \\ I_{n_s} \\ I_I \\ I_O \\ I_{BG} \\ I_{BD} \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \tag{5.7}$$

where

$$Y' = \begin{bmatrix} y_{n_g} & y_{n_d} & (y_{n_s} + y_{n_g} + y_{n_d}) & y_I & y_O & y_{BG} & y_{BD} & y_{n_1} & \cdot & \cdot & y_{n_b} \end{bmatrix}. \tag{5.8}$$

Notice that

$$V_1 = V_{n_g} - V_{n_s}$$

and

$$V_D = V_{n_d} - V_{n_s}.$$

We have

$$Y' \begin{bmatrix} V_1 \\ V_D \\ V_{n_s} \\ V_I \\ V_O \\ V_{BG} \\ V_{BD} \\ V_{n_1} \\ \cdot \\ \cdot \\ V_{n_b} \end{bmatrix} = \begin{bmatrix} I_{n_g} \\ I_{n_d} \\ I_{n_s} \\ I_I \\ I_O \\ I_{BG} \\ I_{BD} \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}. \tag{5.9}$$

Second, we express Y' in its row vector form

$$Y' = \begin{bmatrix} y_1' \\ y_D' \\ y_{n_s}' \\ y_I' \\ y_O' \\ y_{BG}' \\ y_{BD}' \\ y_{n_1}' \\ \cdot \\ \cdot \\ y_{n_b}' \end{bmatrix}, \tag{5.10}$$

where each entry is a row vector. If we add the first and second rows to the third row in (5.9), then, we have

$$\begin{bmatrix} y_1' \\ y_D' \\ (y_{n_s}' + y_1' + y_D') \\ y_I' \\ y_O' \\ y_{BD}' \\ y_{BG}' \\ y_{n_1}' \\ \cdot \\ \cdot \\ \cdot \\ y_{n_b}' \end{bmatrix} \begin{bmatrix} V_1 \\ V_D \\ V_{n_s} \\ V_I \\ V_O \\ V_{BG} \\ V_{BD} \\ V_{n_1} \\ \cdot \\ \cdot \\ \cdot \\ V_{n_b} \end{bmatrix} = \begin{bmatrix} I_{n_g} \\ I_{n_d} \\ 0 \\ I_I \\ I_O \\ I_{BG} \\ I_{BD} \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \tag{5.11}$$

because, from KCL,

$$I_{n_g} + I_{n_d} + I_{n_s} = 0.$$

Now, we suppress internal nodes, from n_1 to n_b, and the node n_s from (5.11). Then, the following equation can be obtained:

$$\begin{bmatrix} y_1'' \\ y_D'' \\ y_I'' \\ y_O'' \\ y_{BD}'' \\ y_{BG}'' \end{bmatrix} \begin{bmatrix} V_1 \\ V_D \\ V_I \\ V_O \\ V_{BG} \\ V_{BD} \end{bmatrix} = \begin{bmatrix} I_{n_g} \\ I_{n_d} \\ I_I \\ I_O \\ I_{BG} \\ I_{BD} \end{bmatrix}, \tag{5.12}$$

where the coefficient matrix is written in the row vector form. Notice that at the output port $I_O = 0$. Hence, V_O can be suppressed so that only V_1, V_D, V_I, V_{BG} and V_{BD} are left in the resultant equations

$$Y^{III}\begin{bmatrix} V_1 \\ V_D \\ V_I \\ V_{BG} \\ V_{BD} \end{bmatrix} = \begin{bmatrix} I_{n_g} \\ I_{n_d} \\ I_I \\ I_{BG} \\ I_{BD} \end{bmatrix}. \tag{5.13}$$

Equation (5.13) can be used at different harmonics. For given V_1 and V_D at a certain frequency $h\omega$, $0 \le h \le H$, and the signal input (when $h = 1$) or DC biases (when $h = 0$), from (5.13), the currents going into the linear part from the interfacing port with nonlinear part can be computed as

$$I_{n_k}(h), \quad 0 \le h \le H, \quad k = g, d, \tag{5.14}$$

where the harmonic index h has been included to indicate the harmonic considered. We use I_{n_k} to denote the complex vector containing all harmonics, that is,

$$I_{n_k} = \begin{bmatrix} I_{n_k}(0) \\ I_{n_k}(1) \\ I_{n_k}(2) \\ \cdot \\ \cdot \\ \cdot \\ I_{n_k}(H) \end{bmatrix}, \quad k = g, d. \tag{5.15}$$

5.2.3 Harmonic Balance Equation

HB starts by assuming an initial guess of V, defined in (5.4), that is

$$V = \begin{bmatrix} V_1 \\ V_G \\ V_D \end{bmatrix},$$

where V_k, $k = 1, G, D$, are given in (5.2). The current summations at the three

nodes, n_g, n_d, and n_{R_1}, are

$$F(V)_{n_g} = I_{n_g}(V) - I_B(V) + j\Omega Q_F(V) + I_G(V) + j\Omega Q_G(V), \tag{5.16}$$

$$F(V)_{n_d} = I_{n_d}(V) + I_B(V) - j\Omega Q_F(V) + I_D(V), \tag{5.17}$$

and

$$F(V)_{n_{R_1}} = I_{R_1}(V) - I_G(V) - j\Omega Q_G(V), \tag{5.18}$$

where I_B, I_G, I_D, I_{R_1}, Q_F and Q_G are from (5.3), I_{n_k}, $k = g, d$, are defined in (5.15), and

$$\Omega = \omega \,\mathrm{diag}\{0, 1, 2, \ldots, H\}. \tag{5.19}$$

In these equations, V is included to symbolize those terms which are functions of V. We obtain the HB equation by assembling, from (5.16) to (5.18),

$$F(V) = \begin{bmatrix} F_{n_g}(V) \\ F_{n_d}(V) \\ F_{n_{R_1}}(V) \end{bmatrix}. \tag{5.20}$$

KCL requires

$$F(V) = 0, \tag{5.21}$$

provided that the V are correctly assumed.

The equation in (5.21) is in complex form. We can rewrite it into a set of algebraic real value equations by splitting the real and imaginary parts,

$$\overline{F}(\overline{V}) = 0, \tag{5.22}$$

where the bar stands for a real vector resulting from the complex counterpart, i.e.,

$$\overline{V} = \begin{bmatrix} \mathrm{Re}(V) \\ \mathrm{Im}(V) \end{bmatrix}, \quad \overline{F} = \begin{bmatrix} \mathrm{Re}(F) \\ \mathrm{Im}(F) \end{bmatrix}.$$

The HB method starts by assuming an initial set $\overline{V}$. Then, a Newton-like

method is used

$$\overline{V}_{new} = \overline{V}_{old} - \overline{J}(\overline{V}_{old})^{-1}\overline{F}(\overline{V}_{old}) \tag{5.23}$$

where $\overline{J}(\overline{V}_{old})$ is the Jacobian and evaluated at $\overline{V}_{old}$.

The Jacobian can be constructed either numerically or analytically. In order to obtain faster convergence and higher accuracy, the analytical Jacobian approach is adopted in our yield optimization.

5.2.4 Response Calculation

Consider the output voltage as the desired response. A linear equation in (5.12) can be used to calculate the output

$$y_O'' \begin{bmatrix} V_1 \\ V_D \\ V_I \\ V_O \\ V_{BG} \\ V_{BD} \end{bmatrix} = I_O. \tag{5.24}$$

If we write out the row vector y_O'' and notice that $I_O = 0$, then we have

$$V_O = -\frac{1}{y''_{O,O}} \begin{bmatrix} y''_{O,1} & y''_{O,D} & y''_{O,I} & y''_{O,BG} & y''_{O,BD} \end{bmatrix} \begin{bmatrix} V_1 \\ V_D \\ V_I \\ V_{BG} \\ V_{BD} \end{bmatrix}. \tag{5.25}$$

It is then possible to find $V_O(h)$, $h = 0, 1, 2, \ldots, H$, after solving the HB equation.

5.3 SPECIFICATIONS AND ERRORS FOR NONLINEAR CIRCUIT YIELD OPTIMIZATION

Consider a nonlinear microwave circuit operating under large-signal steady-state periodic conditions. Response functions for such a circuit may involve DC and harmonic components of the output signal. Unlike linear circuits, therefore, design specifications can be imposed at DC and several harmonics. The jth specification can be denoted by

$$S_{uj}(h), \tag{5.26}$$

if it is an upper specification, or

$$S_{lj}(h) \tag{5.27}$$

in the case of lower specifications, where

$$h = \begin{bmatrix} 0 \\ 1 \\ 2 \\ \cdot \\ \cdot \\ \cdot \\ H \end{bmatrix} \tag{5.28}$$

is the harmonic index vector, 0 and H represent DC and the highest harmonics, respectively. A specific circuit response may involve all or some of the (H + 1) spectral components.

We still use ϕ^i to denote circuit outcomes, namely,

$$\phi^i = \phi^0 + \Delta\phi^i, \quad i = 1, 2, \ldots, K. \tag{5.29}$$

The response of each outcome, denoted by

$$R_i(\phi^i, \overline{V}^i, h) \tag{5.30}$$

is calculated after solving the HB equations

$$\overline{F}(\phi^i, \overline{V}^i) = 0, \tag{5.31}$$

where $\overline{V}^i$ comprises the split real and imaginary parts of the state variables in the HB equation for the outcome ϕ^i. The corresponding error function is defined as

$$R_j(\phi^i, \overline{V}^i, h) - S_{uj}(h) \tag{5.32}$$

or as

$$S_{lj}(h) - R_j(\phi^i, \overline{V}^i, h). \tag{5.33}$$

We assemble all errors for the outcome ϕ^i into one vector c^i. If all entries of this vector are nonpositive, the outcome ϕ^i represents an acceptable circuit.

Following the generalized ℓ_1 centering approach described in Chapter 2, the objective function for the one-sided ℓ_1 optimization is

$$\underset{\phi^0}{\text{minimize}} \left\{ U(\phi^0) = \sum_{i \in I} \sum_{j \in J(\phi^i)} \alpha_i c_j(\phi^i) \right\}, \tag{5.34}$$

where $c_j(\phi^i)$ are elements in $c(\phi^i)$, and $J(\phi^0)$, I, and α_i are defined in (2.36), (2.38) and (2.39), respectively.

5.4 EFFICIENT GRADIENT APPROXIMATION FOR YIELD-DRIVEN DESIGN

It is always a difficult problem to obtain the gradients of circuit response functions when powerful gradient-based optimization is used. For yield-driven design of nonlinear circuits, the computational effort for evaluating gradients may be prohibitive since many nonlinear circuit outcomes are involved. How to obtain the gradient information efficiently is of extreme importance. Much research has been

devoted to devise high-speed gradient approximation methods when analytical gradients are nonexistent or very difficult to get. A simple and traditional way to obtain numerical gradients is the perturbation approximate sensitivity technique (*PAST*)(Rizzoli, Lipparini and Marazzi 1983). Bandler, Chen, Daijavad and Madsen (1988) proposed an integrated gradient approximation technique (*IGAT*) for circuit design.

In the following section we will briefly review *IGAT* in the single function case, then extend *IGAT* to yield optimization.

5.4.1 *IGAT* for Nominal Design

Since the application of *IGAT* is not restricted to circuit response functions, let us use $f(\phi)$ to denote a generic function.

Approximating Derivatives by *PAST*

The first-order derivative of $f(\phi)$ w.r.t the kth variable can be estimated by

$$\frac{\partial f(\phi)}{\partial \phi_k} \approx \frac{f(\phi + \Delta\phi_k u_k) - f(\phi)}{\Delta\phi_k}, \tag{5.35}$$

where $\phi + \Delta\phi_k u_k$ denotes the perturbation of the kth variable, $\Delta\phi_k$ is the perturbation length and u_k is a column vector which has 1 in the kth position and zeros elsewhere. An approximation to the gradient, $\nabla f(\phi)$, can be obtained by perturbing all variables one at a time.

Approximating Derivatives by IGAT

To start the process, *PAST* is used as in Bandler, Chen, Daijavad and Madsen (1988) to calculate the approximate gradient.

The Broyden update generates the new approximate gradient from the previous gradient, namely,

$$\nabla f(\phi_{new}) = \nabla f(\phi_{old}) + \frac{f(\phi_{new}) - f(\phi_{old}) - (\nabla f(\phi_{old}))^T \Delta\phi}{\Delta\phi^T \Delta\phi} \Delta\phi , \quad (5.36)$$

where ϕ_{old} and ϕ_{new} are two different points and $\Delta\phi = \phi_{new} - \phi_{old}$. If ϕ_{old} and ϕ_{new} are iterates of optimization, $f(\phi_{old})$ and $f(\phi_{new})$ need to be evaluated anyway. Thus, the updated gradient can be obtained without additional function evaluations (circuit simulations).

To overcome a particular deficiency of the Broyden update, after a few updates, a special iteration of Powell generates a special step $\Delta\phi$ to guarantee strictly linearly independent directions. After a number of optimization iterations, we may also apply *PAST* to maintain the accuracy of the approximate gradients at a desirable level.

5.4.2 IGAT for Yield Optimization

Notice that $\frac{\partial f(\phi^i)}{\partial \phi_k^i}$ and $\frac{\partial f(\phi^i)}{\partial \phi_k^0}$ are different. The latter is of interest, because ϕ^0 is considered variable in our yield-driven design.

Thus, all perturbations are made to ϕ^0 in the initialization and reinitialization steps using *PAST* (Bandler, Chen, Daijavad and Madsen 1988). When ϕ^0 is perturbed

to $\phi^0 + \Delta\phi_k^0 u_k$, denoted by $\phi_{k,pert}^0$ for short, outcomes should be regenerated from $\phi_{k,pert}^0$ in order to get perturbed circuit responses. These outcomes are denoted by $\phi_{k,pert}^i$. Notice the very same procedure to generate ϕ^i should be repeated to generate $\phi_{k,pert}^i$. It means that the initial status of the random number generation, usually, a set of initial values called seeds, is stored.

From Bandler, Chen, Daijavad and Madsen (1988), the approximate derivative of the response $f(\phi^i)$ is defined as

$$\frac{\partial f(\phi^i)}{\partial \phi_k^0} \approx \frac{f(\phi_{k,pert}^i) - f(\phi^i)}{\Delta\phi_k^0}. \tag{5.37}$$

When the Broyden update or the special iteration of Powell are used, $\Delta\phi^0$ is computed from ϕ_{old}^0 and ϕ_{new}^0 as generated either by the optimization or by the special iteration. Outcomes ϕ_{old}^i and ϕ_{new}^i are outcomes generated from ϕ_{old}^0 and ϕ_{new}^0, respectively. The gradient of the response $f(\phi^i)$ w.r.t. ϕ^0 can be updated as

$$\nabla f(\phi_{new}^i) = \nabla f(\phi_{old}^i) + \frac{f(\phi_{new}^i) - f(\phi_{old}^i) - (\nabla f(\phi_{old}^i))^T \Delta\phi^0}{(\Delta\phi^0)^T \Delta\phi^0} \Delta\phi^0. \tag{5.38}$$

In circuit simulation, there are usually several response levels involved. Suppose the response of interest, on which the design specification is imposed, is the power gain. In the circuit simulation, the power gain is calculated from the output power which, in turn, is calculated from the output voltage. This implies three different response levels. *IGAT* can be applied at any response level.

5.5 NONLINEAR FET STATISTICAL MODELS AND STATISTICAL OUTCOME GENERATION

Purviance, Criss and Monteith (1988) treated the statistical characterization of

small-signal FET models. Our proposed yield optimization requires statistically described large-signal FET models. Our large-signal FET statistical model includes an intrinsic large-signal FET model modified from the Materka and Kacprzak model (Materka and Kacprzak 1985, Microwave Harmonica 1987), statistical distributions and correlations of parameters. The multidimensional normal distribution is assumed for all FET intrinsic and extrinsic parameters. The means and standard deviations are listed in Table 5.1. The correlations between parameters are assumed according to the results published by Purviance, Criss and Monteith (1988). Certain modifications have been made to make the correlations for the large-signal FET model to be consistent with those for the small-signal FET model dealt with in Purviance, Criss and Monteith (1988). The correlation coefficients are given in Table 5.2.

We use a random number generator capable of generating statistical outcomes from the independent and multidimensional correlated normal distributions and from uniform distributions.

Parameters of the nonlinear large-signal models have certain physical limits. A normal distribution random generator may generate outcomes far beyond these limits. These outcomes may cause serious problems either in HB simulation, such as very slow convergence and divergence, or in optimization. Such outcomes must be carefully detected and eliminated.

5.6 YIELD OPTIMIZATION OF A FREQUENCY DOUBLER

5.6.1 Description of the Design

A FET frequency doubler shown in Fig. 5.4 is considered. It consists of a common-source FET with a lumped input matching network and a microstrip output

TABLE 5.1

ASSUMED STATISTICAL DISTRIBUTIONS
FOR THE FET PARAMETERS

FET Parameter	Nominal Value	Standard Deviation (%)	FET Parameter	Nominal Value	Standard Deviation (%)
L_G(nH)	0.16	5	S_I	0.676×10^{-1}	0.65
RD(Ω)	2.153	3	K_G	1.1	0.65
L_S(nH)	0.07	5	τ(pS)	7.0	6
R_S(Ω)	1.144	5	S_S	1.666×10^{-3}	0.65
R_{DE}(Ω)	440	14	I_{G0}(A)	0.713×10^{-5}	3
C_{DE}(pF)	1.15	3	α_G	38.46	3
C_{DS}(pF)	0.12	4.5	I_{B0}(A)	-0.713×10^{-5}	3
I_{DSS}(A)	6.0×10^{-2}	5	α_B	-38.46	3
V_{p0}(V)	-1.906	0.65	R_{10}(Ω)	3.5	8
γ	-15×10^{-2}	0.65	C_{10}(pF)	0.42	4.16
E	1.8	0.65	C_{F0}(pF)	0.02	6.64

The following parameters are considered as deterministic: $K_E = 0.0$, $K_R = 1.111$, $K_1 = 1.282$, $C_{1S} = 0.0$, and $K_F = 1.282$.

For definitions of the FET parameters, see Section 5.2.1.

TABLE 5.2

FET MODEL PARAMETER CORRELATIONS

	L_G	R_S	L_S	R_{DE}	C_{DS}	g_m	τ	R_{IN}	C_{GS}	C_{GD}
L_G	1.00	-0.16	0.11	-0.22	-0.20	0.15	0.06	0.15	0.25	0.04
R_S	-0.16	1.00	-0.28	0.02	0.06	-0.09	-0.16	0.12	-0.24	0.26
L_S	0.11	-0.28	1.00	0.11	-0.26	0.53	0.41	-0.52	0.78	-0.12
R_{DE}	-0.22	0.02	0.11	1.00	-0.44	0.03	0.04	-0.54	0.02	-0.14
C_{DS}	-0.20	0.06	-0.26	-0.44	1.00	-0.13	-0.14	0.23	-0.24	-0.04
g_m	0.15	-0.09	0.53	0.03	-0.13	1.00	-0.08	-0.26	0.78	0.38
τ	0.06	-0.16	0.41	0.04	-0.14	-0.08	1.00	-0.19	0.27	-0.46
R_{IN}	0.15	0.12	-0.52	-0.54	0.23	-0.26	-0.19	1.00	-0.35	0.05
C_{GS}	0.25	-0.24	0.78	0.02	-0.24	0.78	0.27	-0.35	1.00	0.15
C_{GD}	0.04	0.26	-0.12	—0.14	-0.04	0.38	-0.46	0.05	0.15	1.00

Certain modifications have been made to adjust these small-signal parameter correlations (Purviance, Criss and Monteith 1988) to be consistent with the large-signal FET model.

matching and filter section. The optimization variables include the input inductance L_1 and the microstrip lengths l_1 and l_2. Two bias voltages V_{BG} and V_{BD} and the driving power level P_{IN} are also considered as optimization variables. The fundamental frequency is 5GHz. Responses of interest are the conversion gain and spectral purity, defined by

$$\text{conversion gain} = 10\log\frac{\text{power of the second harmonic at the output port}}{\text{power of the fundamental frequency at the input port}}$$

and

$$\text{spectral purity} = 10\log\frac{\text{power of the second harmonic at the output port}}{\text{total power of all other harmonics at the output port}}$$

respectively. The specifications for the conversion gain and spectral purity are 2.5dB and 20dB, respectively. They are both lower specifications.

The large-signal FET statistical model includes an intrinsic large-signal FET model modified from the Materka and Kacprzak model (Materka and Kacprzak 1985), statistical distributions and correlations of parameters. The multidimensional normal distribution is assumed for all FET intrinsic and extrinsic parameters. The means and standard deviations are listed in Table 5.1. The correlation coefficients are given in Table 5.2. Uniform distributions with fixed tolerances of 3% are assumed for P_{IN}, V_{BG}, V_{BD}, L_1, l_1 and l_2. The nominal values for nonoptimizable variables are: L_2 = 15nH, L_3 = 15nH, C_1 = 20pF, C_2 = 20pF, $w_1 = 0.1\text{x}10^{-3}$m, $w_2 = 0.635\text{x}10^{-3}$m, R_{LOAD} = 50Ω, R_{INPUT} = 50Ω, R_{BG} = 10Ω, and R_{BD} = 10Ω. Finally, the uniform distributions with fixed tolerances of 5% are assumed for L_2, L_3, C_1, C_2, w_1 and w_2. The random number generator used is capable of generating outcomes from the independent and multidimensional correlated normal distributions and from uniform distributions.

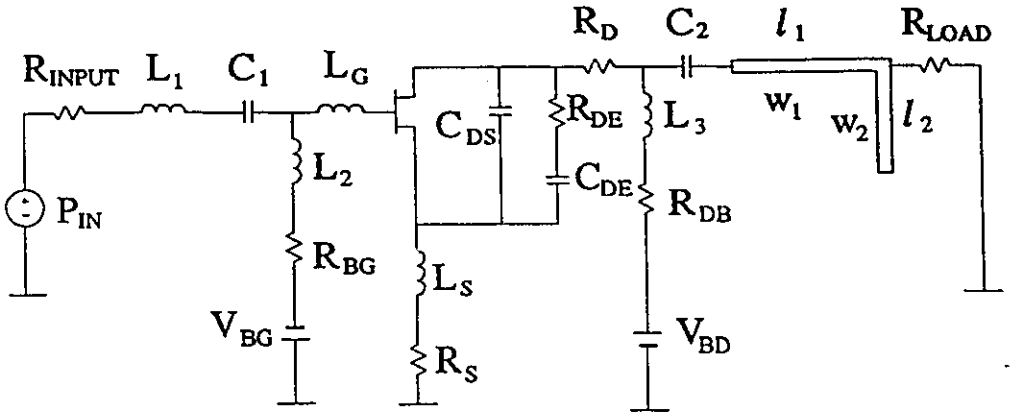

Fig. 5.4 Circuit diagram of the FET microwave frequency doubler.

5.6.2 Design Procedure

In our program, Implementation II of the generalized ℓ_1 centering approach of (2.43) is used. In more detail, the error functions resulting from the simulated conversion gain and spectral purity are calculated, then these error functions with their multipliers defined in (2.43) and (2.39) are fed into the one-sided ℓ_1 optimization. *IGAT* calculates approximate sensitivities of the conversion gain and spectral purity. The computer used is the Multiflow Trace 14/300[1].

The starting point for yield optimization is the solution of the minimax nominal design w.r.t. the same specifications, using the same six design variables. At this point, the estimated yield based on 500 outcomes is 39.6%.

We conduct a design using *IGAT* gradient calculation. Computational details are given in Table 5.3. The design has two consecutive phases, that is, the starting point for the second phase is the solution of the first phase. The second phase is to reoptimize the first solution with updated α_i.

5.6.3 Results and Discussions

Using *IGAT*, the first phase reaches 71% yield and the second phase confirms that the solution of the first phase has been optimized in terms of the estimated yield. The two phases use 61 optimization iterations and 184 function evaluations.

Figs. 5.5, 5.6, 5.7 and 5.8 show histograms of the conversion gain and the spectral purity, respectively. 500 outcomes are used to calculate both distributions.

---

[1]This machine has the performance of 10.04MFLOPS in a test of solving 100 linear equations. The routines used are the standard LINPACK routines. Full precision (64 bits) arithmetic was used. The VAX 11/780 computer has the performance of 0.14MFLOPS for the same test.

TABLE 5.3

YIELD OPTIMIZATION OF THE FREQUENCY DOUBLER USING *IGAT*

Parameter	Starting Point	Nominal Design	Solution I	Solution II
P_{IN}(W)	2.0×10^{-3}	2.49048×10^{-3}	1.98488×10^{-3}	1.92366×10^{-3}
V_{BG}(V)	-1.9	-1.70329	-1.93468	-1.92542
V_{BD}(V)	5.0	6.50000	6.50000	6.50000
L_1(nH)	5.0	5.29066	5.68905	5.633822
l_1(m)	1.0×10^{-3}	1.77190×10^{-3}	1.73378×10^{-3}	1.73740×10^{-3}
l_2(m)	5.0×10^{-3}	5.73087×10^{-3}	5.75011×10^{-3}	5.74907×10^{-3}
Yield		39.6%	71.0%	71.0%
No. of Optimization Iterations			23	38
No. of Circuit Simulations			4450	4750
CPU(Multiflow Trace 14/300)			18.6min	19.1min

The yield is estimated from 500 outcomes.

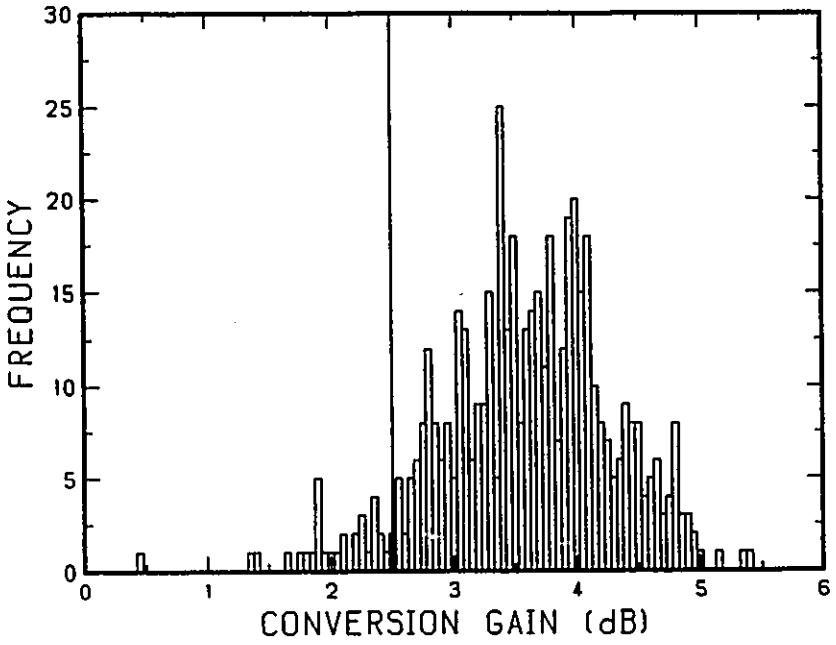

Fig. 5.5 Histogram of conversion gains of the frequency doubler at the starting point. 500 outcomes are used.

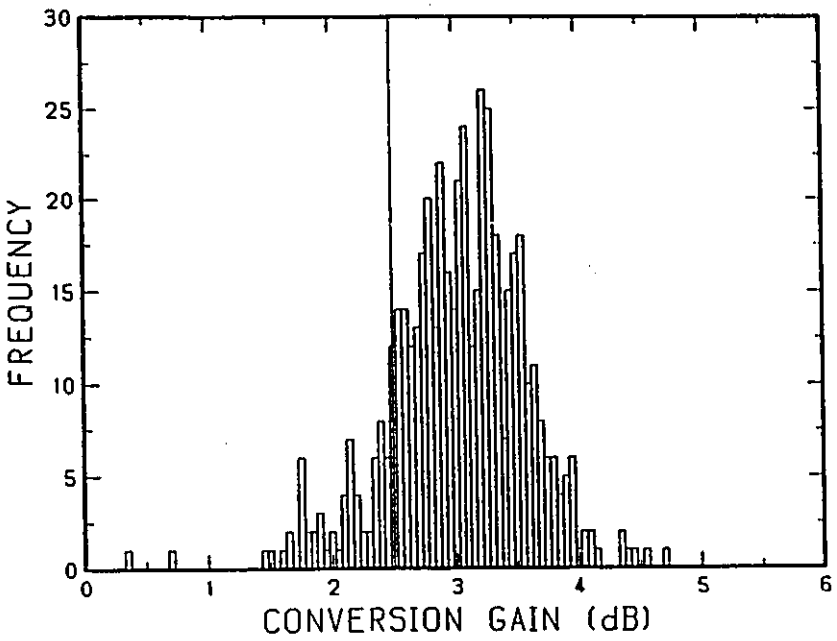

Fig. 5.6 Histogram of conversion gains of the frequency doubler at the solution of yield optimization using *IGAT*. 500 outcomes are used.

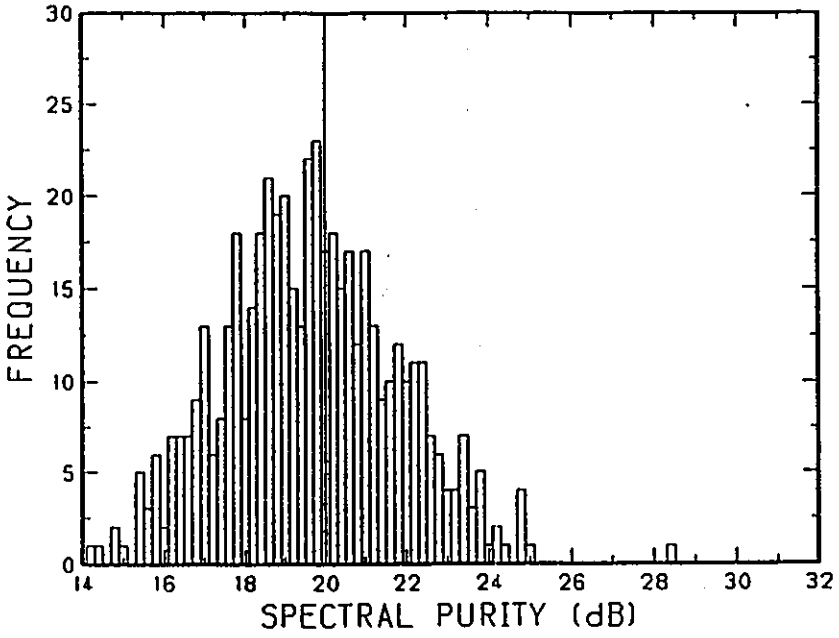

Fig. 5.7 Histogram of spectral purities of the frequency doubler at the starting point. 500 outcomes are used.

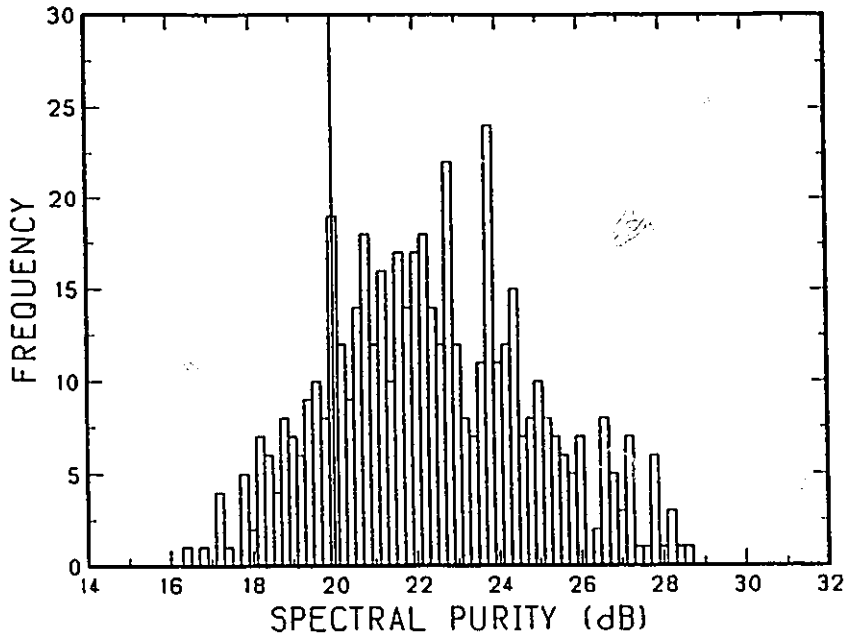

Fig. 5.8 Histogram of spectral purities of the frequency doubler at the solution of yield optimization using *IGAT*. 500 outcomes are used.

Fig. 5.5 is the conversion gain distribution before yield optimization. The histogram in Fig. 5.6 is based on the final solution. Fig. 5.7 is the spectral purity distribution before yield optimization. The histogram in Fig. 5.8 is based on the final solution. The improvement in spectral purity is very clearly illustrated by the histograms in Fig. 5.8. Before yield optimization, the center of the distribution is close to the design specification of 20dB, indicating that many outcomes are unacceptable. After yield optimization, the center of the distribution is shifted to the right-hand side of the specification. Most outcomes then satisfy the specification.

5.7 CONCLUDING REMARKS

We have presented a comprehensive approach to yield-driven design of nonlinear microwave circuits operating within the HB simulation environment. This had been the first convincing demonstration of yield optimization of statistically characterized nonlinear microwave circuits. Our success is due to the sophisticated combination of the following advanced techniques: efficient HB simulation using exact Jacobians, powerful one-sided ℓ_1 design centering, effective and robust gradient approximation, and flexibility of statistical handling to allow different kinds of nonlinear device parameter statistics. In our design example, the large-signal FET statistical model is fully facilitated. Comprehensive numerical experiments directed at yield-driven optimization of a FET frequency doubler verify our approach.

6

FAST GRADIENT BASED NONLINEAR CIRCUIT STATISTICAL DESIGN

6.1 INTRODUCTION

This chapter addresses fast gradient calculation for nonlinear circuit yield-driven design. Statistical design of practical nonlinear microwave circuits is a challenge. One serious inherent difficulty is the prohibitive computational cost: many circuits have to be simulated repeatedly and each circuit simulation involves CPU intensive iterations to solve harmonic balance equations. Furthermore, gradient-based optimization requires effort to estimate the gradients of the error functions. Therefore, an effective and efficient approach to gradient calculation is of the utmost importance.

The conventional *P*erturbation *A*pproximate *S*ensitivity *T*echnique (*PAST*) is conceptually simple. Since *PAST* needs to perturb all variables one at a time, the computational effort involved grows in proportion to the number of variables. Rizzoli, Lipparini and Marazzi (1983) used this method in their single-loop approach to nominal circuit design. In yield optimization, however, *PAST* becomes extremely inefficient because of the large number of circuit outcomes to be dealt with.

The *E*xact *A*djoint *S*ensitivity *T*echnique (*EAST*) has been recently developed by Bandler, Zhang and Biernacki (1988a, 1988b) for the harmonic balance technique. In contrast to *PAST*, *EAST* involves solving a set of linear equations whose coefficient matrix is available after circuit simulation. The solution of a single adjoint system is sufficient for the calculation of sensitivities with respect to all variables. No

perturbation or iterative simulations are required. *EAST* enjoys high computational efficiency, but is very difficult to implement.

We introduce a powerful approach to gradient calculation, the *F*easible *A*djoint *S*ensitivity *T*echnique (*FAST*), by Bandler, Zhang and Biernacki (1989) and Bandler, Zhang, Song and Biernacki (1990). Motivated by the potential impact of the adjoint sensitivity approach on general purpose CAD programs we have studied its implementational aspects. *FAST* combines the efficiency and accuracy of the adjoint sensitivity technique with the simplicity of the perturbation technique. *FAST* is demonstrated to be an implementable, high-speed gradient calculation technique. *FAST* retains most of the efficiency and accuracy of *EAST* while accommodating the simplicity of *PAST*. *FAST*, linking state-of-the-art optimization and efficient harmonic balance simulation, is the key to making our approach to nonlinear microwave circuit design the most powerful available.

First, we describe a general formulation for the adjoint sensitivity analysis technique. Then, this formulation is applied to develop *FAST* for the nominal design case. *FAST* is extended to a form applicable to statistical design. Detailed comparisons among *FAST*, *IGAT*, *PAST*, and *EAST* are given.

FAST is used in yield optimization of a microwave frequency doubler. The performance yield is increased from 40% to 70%. This example is the same circuit used in the last chapter so that the comparison between *IGAT* and *FAST* can be drawn. *FAST* exhibits superior efficiency.

6.2 FEASIBLE ADJOINT SENSITIVITY TECHNIQUE

6.2.1 Generic Formulas for Adjoint Sensitivity Analysis

We derive a set of generic formulas to calculate sensitivities of a response in the harmonic balance simulation environment. The sensitivity of a response with respect to one variable, ϕ_k,

$$\frac{\partial R_j(\phi, \overline{V}, h)}{\partial \phi_k} \tag{6.1}$$

should be computed at the solution of the harmonic balance equation, that is, this sensitivity calculation is performed subject to the constraints of

$$\overline{F}(\phi, \overline{V}) = 0.$$

First, (6.1) is found to be

$$\frac{\partial R_j(\phi, \overline{V}, h)}{\partial \phi_k}\Big|_{\overline{F}=0} = \frac{\partial R_j(\phi, \overline{V}, h)}{\partial \phi_k} + \left[\frac{\partial R_j(\phi, \overline{V}, h)}{\partial \overline{V}}\right]^T \frac{\partial \overline{V}}{\partial \phi_k}. \tag{6.2}$$

Then, we differentiate both sides of the harmonic balance equation with respect to ϕ_k,

$$\frac{\partial \overline{F}(\phi, \overline{V})}{\partial \phi_k} + \left[\frac{\partial \overline{F}(\phi, \overline{V})}{\partial \overline{V}}\right]^T \frac{\partial \overline{V}}{\partial \phi_k} = 0. \tag{6.3}$$

Now, from (6.3), we can find that

$$\begin{aligned} \frac{\partial \overline{V}}{\partial \phi_k} &= -\left[\left(\frac{\partial \overline{F}}{\partial \overline{V}}\right)^T\right]^{-1} \frac{\partial \overline{F}}{\partial \phi_k} \\ &= -\overline{J}^{-1} \frac{\partial \overline{F}}{\partial \phi_k}, \end{aligned} \tag{6.4}$$

where $\overline{J}$ is the Jacobian matrix defined in (5.23).

By substituting (6.4) into it, (6.2) can be rewritten as:

$$\frac{\partial R_j(\phi, \overline{V}, h)}{\partial \phi_k}\Big|_{\overline{F}=0} = \frac{\partial R_j(\phi, \overline{V}, h)}{\partial \phi_k} - \left[\frac{\partial R_j(\phi, \overline{V}, h)}{\partial \overline{V}}\right]^T J^{-1}\frac{\partial \overline{F}}{\partial \phi_k}. \qquad (6.5)$$

Now the three terms on the right-hand side of (6.5) should be available before computing the sensitivity. For different types of responses and variables, those terms can become very involved. (It is a headache to implement this approach in a general-purpose program while various responses and user-defined responses are allowed.) Bandler, Zhang and Biernacki (1988a, 1988b) have proposed *EAST* to calculate sensitivities. The sensitivity expressions for various elements have been derived and listed (Bandler, Zhang and Biernacki 1988b).

6.2.2 *FAST* for Nominal Design

To simplify the derivation and to embody the procedure, we consider the circuit in Fig. 5.1 in the last chapter. The response of interest is the voltage at the output port in the circuit shown in Fig. 5.2. The output voltage at a certain harmonic can be calculated from (5.25). We may want to include $V_G(h)$ into (5.25),

$$V_O(h) = -\frac{\left[y''_{O,1}(h) \quad 0 \quad y''_{O,D}(h) \quad y''_{O,I}(h) \quad y''_{O,BG}(h) \quad y''_{O,BD}(h)\right]}{y''_{O,O}(h)} \begin{bmatrix} V_1(h) \\ V_G(h) \\ V_D(h) \\ V_I(h) \\ V_{BG}(h) \\ V_{BD}(h) \end{bmatrix}. \qquad (6.6)$$

We define a vector for the output voltage spectra,

$$V_O \triangleq \begin{bmatrix} V_O(0) \\ V_O(1) \\ \cdot \\ \cdot \\ \cdot \\ V_O(H) \end{bmatrix}.$$

We also define a vector containing the input and DC bias voltages

$$V_S \triangleq \begin{bmatrix} V_I(1) \\ V_{BG}(0) \\ V_{BD}(0) \end{bmatrix}.$$

Now, we can relate the output voltage at all harmonics to the harmonic balance solution V, input excitation and DC bias supplies

$$Y_{O,O} V_O = \begin{bmatrix} Y_{O,F} & Y_{O,S} \end{bmatrix} \begin{bmatrix} V \\ V_S \end{bmatrix}, \tag{6.8}$$

where

$$Y_{O,O} \triangleq \text{diag}\{ y''_{O,O}(0)\ y''_{O,O}(1) \cdots y''_{O,O}(H) \}\ , \tag{6.9}$$

$$Y_{O,F} \triangleq -\left[\begin{array}{cccccc|c|cccccc} y''_{O,1}(0) & 0 & \cdot & \cdot & \cdot & \cdot & & y''_{O,D}(0) & 0 & \cdot & \cdot & \cdot & \cdot \\ 0 & y''_{O,1}(1) & 0 & \cdot & \cdot & \cdot & & 0 & y''_{O,D}(1) & 0 & \cdot & \cdot & \cdot \\ & & \cdot & & & & 0 & & & \cdot & & & \\ & & & \cdot & & & & & & & \cdot & & \\ & & & & \cdot & & & & & & & \cdot & \\ \cdot & \cdot & \cdot & \cdot & 0 & y''_{O,1}(H) & & \cdot & \cdot & \cdot & \cdot & 0 & y''_{O,D}(H) \end{array}\right]$$

(6.10)

and

$$Y_{O,S} \triangleq -\begin{bmatrix} 0 & y''_{O,BG}(0) & y''_{O,BD}(0) \\ y''_{O,I}(1) & 0 & 0 \\ 0 & 0 & 0 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 0 & 0 & 0 \end{bmatrix}. \tag{6.11}$$

From (6.8), we have

$$V_O = Y_{O,O}^{-1}\begin{bmatrix} Y_{O,F} & Y_{O,S} \end{bmatrix}\begin{bmatrix} V \\ V_S \end{bmatrix}, \tag{6.12a}$$

or

$$V_O = \begin{bmatrix} A & B \end{bmatrix}\begin{bmatrix} V \\ Y_S \end{bmatrix}, \tag{6.12b}$$

where matrices A and B are appropriately defined.

By splitting the real and imaginary parts of (6.12), the real-valued version can be found as

$$\overline{V}_O \triangleq \begin{bmatrix} \mathrm{Re}(V_O) \\ \mathrm{Im}(V_O) \end{bmatrix} = \begin{bmatrix} \mathrm{Re}(A) & -\mathrm{Im}(A) & \mathrm{Re}(B) & -\mathrm{Im}(B) \\ \mathrm{Im}(A) & \mathrm{Re}(A) & \mathrm{Im}(B) & \mathrm{Re}(B) \end{bmatrix}\begin{bmatrix} \mathrm{Re}(V) \\ \mathrm{Im}(V) \\ \mathrm{Re}(V_S) \\ \mathrm{Im}(V_S) \end{bmatrix}.$$

Now, we define the following

$$\begin{bmatrix} \overline{A} & \overline{B} \end{bmatrix}\begin{bmatrix} \overline{V} \\ \overline{V}_S \end{bmatrix} \triangleq \begin{bmatrix} \mathrm{Re}(A) & -\mathrm{Im}(A) & \mathrm{Re}(B) & -\mathrm{Im}(B) \\ \mathrm{Im}(A) & \mathrm{Re}(A) & \mathrm{Im}(B) & \mathrm{Re}(B) \end{bmatrix}\begin{bmatrix} \mathrm{Re}(V) \\ \mathrm{Im}(V) \\ \mathrm{Re}(V_S) \\ \mathrm{Im}(V_S) \end{bmatrix}, \tag{6.13}$$

where $\overline{V}_S$ denotes the split real and imaginary parts in the spectra of excitation voltages, and $\overline{V}$ is actually the solution to the harmonic balance equations (5.22). Then,

$$\overline{V}_O = [\overline{A}\ \ \overline{B}] \begin{bmatrix} \overline{V} \\ \overline{V}_S \end{bmatrix} \Delta\, \overline{C} \begin{bmatrix} \overline{V} \\ \overline{V}_S \end{bmatrix} \tag{6.14}$$

has a linear transfer matrix linking the output voltage with $\overline{V}_S$ and $\overline{V}$. The coefficients of $\overline{C}$ are functions of element parameters in the linear subnetwork. $\overline{V}$ are dependent not only upon element parameters in both linear and nonlinear subnetworks, but also the excitation and DC biases. Obviously, $\overline{V}_S$ is determined by the excitation voltage and DC bias voltages. We may want to change $\overline{V}_S$ to improve the circuit performance.

To make the formulation concise, we only consider the derivative of

$$V_O^R(h) = \mathrm{Re}\{V_O(h)\}$$

with respect to ϕ_k. The corresponding row in (6.14) gives

$$V_O^R(h) = \begin{bmatrix} a^T & b^T \end{bmatrix} \begin{bmatrix} \overline{V} \\ \overline{V}_S \end{bmatrix} = c^T \begin{bmatrix} \overline{V} \\ \overline{V}_S \end{bmatrix}. \tag{6.15}$$

From this equation, the approximate derivative of V_O^R w.r.t. ϕ_k can be calculated as

$$\frac{\Delta V_O^R}{\Delta \phi_k} \approx \frac{\Delta c^T}{\Delta \phi_k} \begin{bmatrix} \overline{V} \\ \overline{V}_S \end{bmatrix} + a^T \frac{\Delta \overline{V}}{\Delta \phi_k} + b^T \frac{\Delta \overline{V}_S}{\Delta \phi_k} \tag{6.16}$$

by perturbing ϕ to $\phi + \Delta\phi_k u_k$.

Let $\overset{\wedge}{\overline{V}}$ be the adjoint voltages obtained by solving the linear equation

$$\overline{J}^T \overset{\wedge}{\overline{V}} = a, \tag{6.17}$$

where $\overline{J}$ is the Jacobian of $\overline{F}$ with respective to $\overline{V}$ at the harmonic balance solution. Substituting (6.17) into the third term on the right hand side of (6.16) gives the

$$\overset{\wedge}{\overline{V}}^T \overline{J} \Delta \overline{V}.$$

From (6.4), we have

$$-\overline{J}\,\Delta\overline{V} \approx \Delta\overline{F}.$$

Now, we can express

$$\Delta V_O^R \approx \left[\, c^T(\phi + \Delta\phi u_k) - c^T \,\right] \begin{bmatrix} \overline{V} \\ \overline{V}_S \end{bmatrix} + b^T\left[\overline{V}_S(\phi + \Delta\phi u_k) - \overline{V}_S\right] - \overline{\overline{V}}^T \Delta\overline{F}. \tag{6.18}$$

The incremental term $\Delta\overline{F}$ can be approximated by

$$\Delta\overline{F} \approx \overline{F}(\phi + \Delta\phi_k u_k, \overline{V}) \tag{6.19}$$

for a small $\Delta\phi_k$.

Equation (6.18) is a special case of (6.5).

Considering the different elements, (6.18) can be further expressed as

$$\Delta V_O^R \approx \begin{cases} \left[c^T(\phi + \Delta\phi_k u_k) - c^T\right] \begin{bmatrix} \overline{V} \\ \overline{V}_S \end{bmatrix} - \overline{\overline{V}}^T \overline{F}(\phi + \Delta\phi_k u_k, \overline{V}), \\ \qquad \text{if } \phi_k \in \text{linear subnetwork;} \\ b^T\left[\overline{V}_S(\phi + \Delta\phi_k u_k) - \overline{V}_S\right] - \overline{\overline{V}}^T \overline{F}(\phi + \Delta\phi_k u_k, \overline{V}), \\ \qquad \text{if } \phi_k \in \text{sources;} \\ -\overline{\overline{V}}^T \overline{F}(\phi + \Delta\phi_k u_k, \overline{V}), \\ \qquad \text{if } \phi_k \in \text{nonlinear subnetwork.} \end{cases} \tag{6.20}$$

This formula is much easier to implement than the corresponding formula for *EAST* (Bandler, Zhang and Biernacki 1988b). The function $\overline{F}(\phi + \Delta\phi_k u_k, \overline{V})$ is evaluated by perturbation. The effort for solving the linear equations (6.17) is small

since the LU factors of the Jacobian matrix are already available from the final harmonic balance iteration. The terms $\overline{V}$ and $\overline{V}_S$ are also available from the harmonic balance simulation. The perturbed vectors $a(\phi + \Delta\phi_k u_k)$ and $b(\phi + \Delta\phi_k u_k)$ can be easily calculated since they involve the linear subnetwork only. Finally, the perturbed excitations $\overline{V}_S(\phi + \Delta\phi_k u_k)$ can be effortlessly obtained. It is clear that the calculation of all the terms in (6.18) or (6.20) can be readily implemented.

Finally, the approximate sensitivity of output voltage $V_O^R(h)$ with respective to ϕ_k can be computed as

$$\frac{\partial V_O^R(h)}{\partial \phi_k} \approx \frac{\Delta V_O^R(h)}{\Delta \phi_k}. \tag{6.21}$$

6.2.2 *FAST* for Yield Optimization

For the ith outcome, the following terms are defined, corresponding to their counterparts in the nominal circuit case,

$$V_O^R(\phi^i),\ a(\phi^i),\ b(\phi^i),\ c(\phi^i),\ \overline{V}(\phi^i),\ \overline{V}_S(\phi^i). \tag{6.22}$$

The harmonic balance equation and adjoint system are

$$\overline{F}(\phi^i, \overline{V}(\phi^i)) = 0 \tag{6.23}$$

and

$$\left[\overline{J}^T(\phi^i)\right]^{-1} \overline{\Lambda}(\phi^i) = a(\phi^i). \tag{6.24}$$

Similar to *IGAT* perturbations in yield optimization, perturbations used in *FAST* are also made to the nominal values ϕ^0. ϕ^i and $\phi^i_{k,pert}$ are outcomes generated from the unperturbed and perturbed nominal values ϕ^0 and $\phi^0_{k,pert}$, respectively. The increment of the output voltage of the ith outcome due to the perturbation is

calculated by

$$\Delta V_O^R(\phi^i) \approx \left[c^T(\phi^{\,i}_{k,pert}) - c^T(\phi^i)\right]\begin{bmatrix}\overline{V}(\phi^i)\\ \overline{V}_S(\phi^i)\end{bmatrix} + b^T(\phi^i)\left[\overline{V}_S(\phi^{\,i}_{k,pert}) - \overline{V}_S(\phi^i)\right] - \overline{\hat{V}}^T(\phi^i)\Delta\overline{F}\,, \tag{6.25}$$

where

$$\Delta\overline{F} \approx \overline{F}(\phi^i_{k,pert}, \overline{V}(\phi^i)), \tag{6.26}$$

$\overline{V}(\phi^i)$ is the harmonic balance solution of the ith outcome, and $\overline{\hat{V}}(\phi^i)$ is the adjoint system solution corresponding to (6.17).

For different types of elements, a formula similar to (6.20) can be derived.

Although we provided formulas to calculate the sensitivity of the output voltage, the principles behind our approach are applicable to other forms of responses of interest. Since many other responses are functions of the output voltage, their sensitivities can be calculated from the sensitivities of the output voltage. In nonlinear microwave circuit design, the power performance is usually of major interest. The sensitivity information of the output voltage can be translated into that of output power through the chain rule.

6.3 COMPARISONS OF VARIOUS APPROACHES

6.3.1 Implementational Comparisons of *PAST*, *IGAT*, *EAST* and *FAST*

PAST and *IGAT* do not need any modification of the circuit simulator.

PAST is a widely used approach, because it is very easy to implement. However, computational costs may be prohibitive. Suppose there are 10 design

variables in the nonlinear circuit. Using *PAST* to calculate the gradient, one needs to perturb all design variables and to solve the entire nonlinear circuit for each perturbation, i.e., 10 times. The best possible situation for this approach is that all 10 simulations use the same Jacobian and all converge in one iteration. This applies to nominal circuit design. For yield optimization, a large number of statistical outcomes may make *PAST* prohibitive.

The distinct advantage of *IGAT* over *PAST* is that *IGAT* only requires the circuit response function once to update the previously calculated gradient for most optimization iterations. *IGAT* enjoys the simplicity of the perturbation method so that yield optimization can be carried out without modifying the circuit simulator to calculate exact derivatives. *IGAT* is very desirable when the circuit simulator cannot be modified.

Both *EAST* and *FAST* require modification to the circuit simulator.

The generic exact adjoint sensitivity technique (Bandler, Zhang and Biernacki 1988a, 1988b) is accepted by all circuit theoreticians as the most powerful tool. However, to implement it, we have to keep track of all arbitrary locations of variables and to compute branch voltages at all these locations. Microwave software engineers have, to date, found these obstacles insurmountable.

Using *FAST*, we also need to perturb all variables. For a circuit with 10 design variables, instead of completely solving 10 nonlinear circuits, we only evaluate 10 residuals in the form of (6.19) and calculate the perturbed linear subnetwork. The solution of adjoint equation (6.17) can be accomplished by using forward and backward substitutions. In *FAST*, we completely eliminate the need to track variable locations. We only need to identify the output port, which is the simplest step in

adjoint sensitivity theory.

6.3.2 Numerical Comparison of *PAST*, *EAST* and *FAST*

We use a MESFET mixer (Camacho-Penalosa and Aitchison 1987, Bandler, Zhang and Biernacki 1988a, 1988b) to investigate the accuracy and actual time efficiency of *FAST* (Bandler, Zhang and Biernacki 1989). Sensitivities of the mixer conversion gain w.r.t. 26 variables were calculated by the *FAST*, *EAST* and *PAST* approaches, respectively. The variables include all parameters in the linear as well as in the nonlinear part, DC bias, LO power, IF, LO and RF terminations. The results show that the *FAST* sensitivities are almost identical to the exact sensitivities, whereas the sensitivities computed by *PAST* are typically 1 to 2 percent different from their exact values. This fact reveals that *FAST* promises to be much more reliable than *PAST*. The CPU time comparison shows that *FAST* is 3 times slower than *EAST* but 23 time faster than *PAST* for one complete sensitivity analysis of the mixer circuit.

The comparison between *IGAT* and *FAST* will be given in the following section by a statistical design example.

6.4 YIELD OPTIMIZATION OF A FREQUENCY DOUBLER

The same FET frequency doubler shown in Fig. 5.1 in Chapter 5 is considered. We implement *FAST* and conduct a design using *FAST* gradient calculation in the same environment as *IGAT* was implemented. Computational details are given in Table 6.1. The design has two consecutive phases. The first phase uses 19 function evaluations and gradient calculations to give 70.6% yield. The second phase slightly increases the estimated yield to 71%, verifying the solution of the first phase. The

TABLE 6.1

YIELD OPTIMIZATION OF THE FREQUENCY DOUBLER USING *FAST*

Parameter	Starting Point	Nominal Design	Solution I	Solution II
P_{IN}(W)	2.0×10^{-3}	2.49048×10^{-3}	2.02313×10^{-3}	1.94444×10^{-3}
V_{BG}(V)	-1.9	-1.70329	-1.93930	-1.92927
V_{BD}(V)	5.0	6.50000	6.50000	6.50000
L_1(nH)	5.0	5.29066	5.71547	5.63312
l_1(m)	1.0×10^{-3}	1.77190×10^{-3}	1.73531×10^{-3}	1.74046×10^{-3}
l_2(m)	5.0×10^{-3}	5.73087×10^{-3}	5.74965×10^{-3}	5.74956×10^{-3}
Yield		39.6%	70.6%	71.0%
No. of Optimization Iterations			19	29
No. of Circuit Simulations and FAST Analyses			950	1450
CPU(Multiflow Trace 14/300)			7.9min	12.1min

The yield is estimated from 500 outcomes.

efficiency of *FAST* is well demonstrated. To reach the same yield level, the CPU time used by the *FAST* approach is much less than that used by the *IGAT* approach.

Figs. 6.1 and 6.2 show histograms of the conversion gains and the spectral purities, respectively, at the solution of yield optimization. 500 outcomes are used to calculate both distributions. The histograms of Fig. 6.1 and Fig. 6.2 are very similar to those in Fig. 5.6 and Fig. 5.8, respectively. The center of the spectral purity histogram has been moved well above the specification.

6.5 CONCLUDING REMARKS

This chapter has presented the comprehensive formulation of gradient calculation of *FAST*. Combining perturbations, and adjoint analysis techniques, *FAST* has significantly improved computational efficiency as compared with most existing methods. The significant advantages *FAST* over *PAST* and *IGAT* are its unmatched speeds and accuracy, and over *EAST* are its implementational simplicity. *IGAT* is a desirable choice when the circuit simulator cannot be modified. *FAST* is particularly suitable for implementation in general purpose microwave CAD software.

Numerical experiments directed at yield-driven optimization of a FET frequency doubler verify the new gradient calculation approach. The substantial computational advantage of *IGAT* and *FAST* have been observed.

Since most microwave circuit CAD packages currently used in industry utilize the traditional perturbation approach to evaluate gradients, design optimization is severely limited by poor efficiency. Because of its superior properties, the gradient calculation technique based on adjoint sensitivity analysis has attracted much attention from the microwave CAD community (Gilmore and Steer 1991). The author strongly

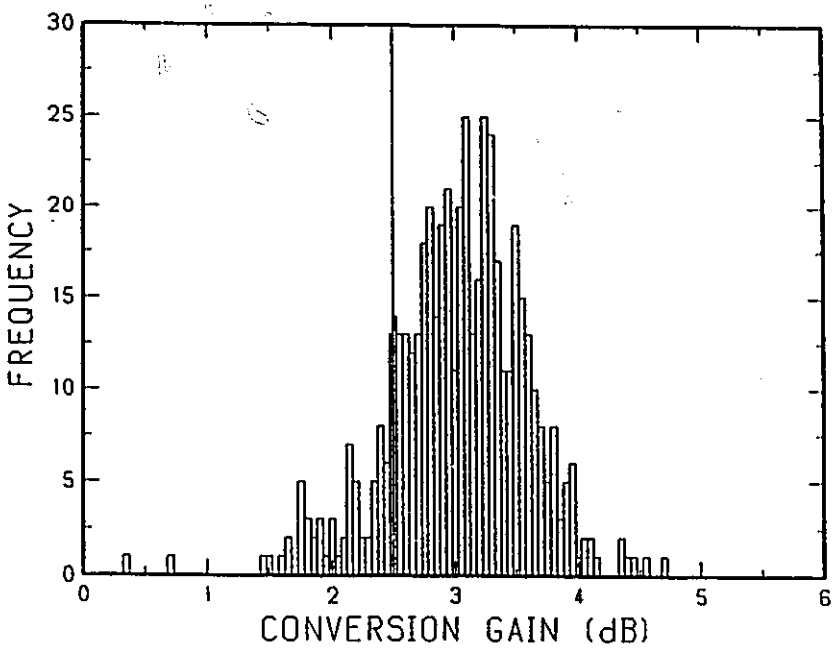

Fig. 6.1 Histogram of conversion gains of the frequency doubler at the solution of yield optimization using *FAST*. 500 outcomes are used.

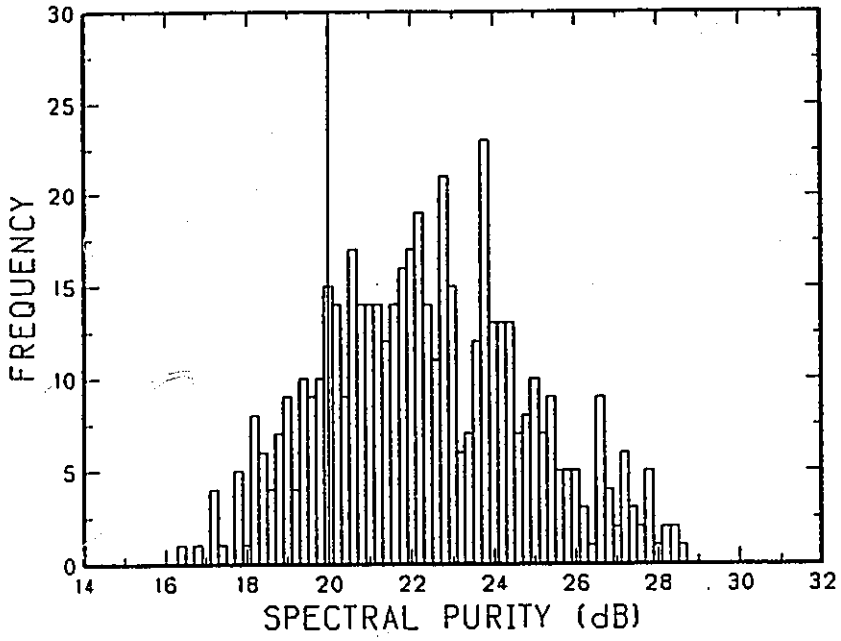

Fig. 6.2 Histogram of spectral purities of the frequency doubler at the solution of yield optimization using *FAST*. 500 outcomes are used.

believes that *FAST* will become an expedient tool to meet the pressing need for efficient microwave nonlinear circuit design. Our success should strongly motivate other developments in nonlinear microwave circuit CAD, such as yield-driven design using physicas based device models and statistical modelling of microwave devices for large-signal applications.

7

CONCLUSIONS

This thesis has addressed yield-driven design of microwave circuits using efficient computer-aided design techniques. Novel approaches to approximations to both circuit responses and gradients have been described. Emphasis has been put on the feasibility of these approaches. This is essential both in dealing with practical large problems and allowing yield-driven design to exist in general-purpose CAD tools. Our approaches offer great reductions in computational requirements, including storage and CPU time. Meanwhile, it is easy to implement them into a new CAD programs or to integrate them with existing programs.

The efficient quadratic modeling presented in Chapter 3 offers substantial savings of CPU time and storage through avoiding repeated expensive simulations and gradient evaluations. This technique can be applied to a broad range of applications where many computationally intensive simulations are required, such as statistical modeling, Monte Carlo simulation, iterative numerical calculation involving very complicated function evaluations, etc. For example, it has been suggested by Rizzoli et al. (1991) that complex components that can only be computed by electromagnetic methods are approximated by our quadratic approximation.

Actual tuning is always associated with tolerances. In Chapter 4, yield-driven design of such circuits is discussed. By taking tuning tolerances into consideration, our yield-driven design will ease the tuning process and the design results present better yields after tuning. For very large scale circuits, yield-driven design using

general-purpose CAD programs on available computer platforms may become unmanageable. Our combined approach presented in Chapter 4 provides an effective means to meet this challenge. Due to our work, the application of supercomputers to yield-driven design task has generated interests in microwave CAD area (Rizzoli et al. 1991)

Our approaches to gradient calculation in the harmonic balance environment, presented in Chapters 5 and 6, are directly applicable to yield-driven design using optimization techniques. They have shown very promising features including high computational efficiency and implementational feasibility. We have illustrated how these methods can be implemented to suit different situations. This work has been highly regarded by other researchers. Purviance and Meehan (1991) stated that the step from linear circuits to nonlinear circuits in yield optimization was first presented to the microwave CAD community by us. These methods have found their way into general-purpose CAD software. We believe that they will play an important role in the new generation microwave circuit CAD software. Reliable and efficient yield-driven techniques will also facilitate and broaden the new CAD research.

Comprehensive testing examples have been presented in this thesis to examine our new approaches. The merits of these approaches have been well demonstrated.

The author has felt, during the course of this work, that the issue of how to improve manufacturability and yield has been realized to be of extreme importance by more and more engineers in the MMIC field. Development of reliable and efficient yield-driven design tools has lagged behind the demand. In order to reduce this gap, the following problems could be considered in future research and development.

(a) In the proposed gradient quadratic approximation scheme of Chapter 3, a circuit response and its gradient are treated as separated functions. However, they are actually related. Therefore, it is suggested that the gradient information, together with response information, should be utilized to build the response model. Then, consistency between the quadratic models of the response and gradient and higher accuracy of the response model are logical consequences. A problem is that the resultant formulas may not be as simple as those in Chapter 3.

(b) Statistical design of large scale circuits is still one of the most challenging tasks in CAD. Approximation techniques will continue to play an extremely important role in the coming generation of CAD. They can be applied at the system response level, subsystem level, and/or device and element level. Besides some traditional approximation techniques, such as quadratic approximation of Low and Director (1989) and spline approximation of Barby, Vlach and Singhal (1988), the response surface method based on experimental design theory (Box, Hunter and Hunter 1978) has be used to guide statistical design of VLSI devices (Alvarez, Abdi, Young, Weed and Herald 1988, McDonald, Maini, Spangler and Weed 1989, and Aoki, Masuda, Shimada, and Sato 1987). Applying this method to microwave circuits and devices would tremendously improve the efficiency of large scale network yield-driven design.

(c) As we know, the ultimate goal of design is to direct the manufacturing

process. It is much more meaningful to use actual geometric sizes and process controlling parameters as design variables. Physics based models of passive and active devices are necessary to link the design procedure and the manufacturing process (Yoshii, Tomizawa and Yokoyama 1983, Snowden and Loret 1987, Khatibzadeh and Trew 1988, and Curtice 1989). Research into using physics based device models in design has been undertaken by Bandler, Zhang and Cai (1990). Simulation of active devices, such as GaAs FETs, is still the computational bottleneck. In the case of yield-driven design with physics based models, computational effort will be even more prohibitive. Approximation to device level simulation should be devised. The table lookup approach (Burns, Newton and Pederson 1983, Jain, Agnew and Nakhla 1983) and the macromodeling approach (Turchetti and Masetti 1983, Casinovi and Sangiovanni-Vincentelli 1991) can be very good candidates to be applied in microwave circuit design. It is better to complete the approximation calculation in the preprocessing phase prior to actual design optimization.

(d) Another urgent problem emerging is statistical modeling of different types of elements and devices. Reliable statistical models are key to determining accuracy of yield estimates and, consequently, of yield-driven design. The more difficult aspect will be the statistical modeling of devices using physics based models. Research reported is mainly focused on digital circuit devices (Cox, Yang, Mahant-Shetti and Chatterjee 1985, Herr and Barnes 1986, and Yu, Kang, Hajj and Trick 1987). One of the first research results related to the statistical modeling of GaAs MESFETs is given by Bandler, Biernacki,

Chen, Song, Ye and Zhang (1991b). Further research is definitely needed to produce efficient, consistent and unique solutions.

(e) Parallel processing and distributed computation will provide us with even greater computational power. Calculations required in yield-driven design have an inherent parallel nature which is highly suitable for parallel processing (Rizzoli et al. 1991). Research into this parallelism should be conducted to efficiently explore available networked computers.

REFERENCES

H.L. Abdel-Malek and J.W. Bandler (1980), "Yield optimization for arbitrary statistical distributions: Part I-theory", *IEEE Trans. Circuits Syst.*, vol. CAS-27, pp. 245-253.

A.R. Alvarez, B.L. Abdi, D.L. Young, H.D. Weed and E.R. Herald (1988), "Application of statistical design and response surface methods to computer-aided VLSI device design", *IEEE Trans. Computer-Aided Design*, vol. 7, pp. 272-288.

K.J. Antreich and R.K. Koblitz (1982),"Design centering by yield prediction", *IEEE Trans. Circuits Syst.*, vol. CAS-29, pp. 88-96.

Y. Aoki, H. Masuda, S. Shimada and S. Sato (1987), "A new design-centering methodology for VLSI device development," *IEEE Trans. Computer-Aided Design*, vol. CAD-6, pp. 452-461.

A.E. Atia (1974), "Computer-aided design of microwave multiplexers", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-22, pp. 332-336.

I. Bahl and P. Bhartia (1988), *Microwave Solid State Circuit Design*. New York: Wiley.

J.W. Bandler (1972), "Optimization of design tolerances using nonlinear programming", *Proc. 6th Princeton Conf. Information Sciences and System* (Princeton, N.J.), pp. 655-659. Also in *Computer-Aided Filter Design*, G. Szentirmai, Ed., New York: IEEE Press, 1973.

J.W. Bandler (1973), "Computer-aided circuit optimization", in *Modern Filter Theory and Design*, G.C. Temes and S.K. Mitra, Eds. New York: Wiley-Interscience.

J.W. Bandler (1974), "Optimization of design tolerances using nonlinear programming", *J. Optimization Theory Appl.*, vol. 14, pp. 99-114.

J.W. Bandler and H.L. Abdel-Malek (1978), "Optimal centering, tolerancing, and yield determination via updated approximations and cuts", *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 853-871.

J.W. Bandler, R.M. Biernacki, S.H. Chen, M.L. Renault, J. Song and Q.J. Zhang (1988), "Yield optimization of large scale microwave circuits", *Proc. 18th European Microwave Conf.* (Stockholm, Sweden), pp. 255-260.

J.W. Bandler, R.M. Biernacki, S.H. Chen, J. Song, S. Ye, and Q. Zhang (1990), "Analytically unified DC/small-signal/large-signal circuit design", *Proc. Third Asia-Pacific Microwave Conf.* (Tokyo, Japan), pp. 453-456.

J.W. Bandler, R.M. Biernacki, S.H. Chen, J. Song, S. Ye, and Q. Zhang (1991a), "Analytically

unified DC/small-signal/large-signal circuit design", *IEEE Trans. Microwave Theory Tech.* vol. 39.

J.W. Bandler, R.M. Biernacki, S.H. Chen, J. Song, S. Ye, and Q. Zhang (1991b), "Statistical modeling of GaAs MESFETs", *IEEE Int. Microwave Symp. Dig.* (Boston, MA).

J.W. Bandler, R.M. Biernacki, S.H. Chen, J. Song, S. Ye, and Q. Zhang (1991c), "Gradient quadratic approximation scheme for yield-driven design", *IEEE Int. Microwave Symp. Dig.* (Boston, MA).

J.W. Bandler and S.H. Chen (1988), "Circuit optimization: the state of the art", *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 424–443.

J.W. Bandler, S.H. Chen and S. Daijavad (1986), "Exact sensitivity analysis for optimization of multi-coupled cavity filters", *Int. J. Circuit Theory and Appl.*, vol. 14, pp. 63–77.

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 (1988), "Efficient optimization with integrated gradient approximations", *IEEE Trans. Microwave Theory Tech.*, vol. 36. pp. 444–454.

J.W. Bandler, S.H. Chen and K. Madsen (1988), "An algorithm for one-sided ℓ_1 optimization with application to circuit design centering", *Proc. IEEE Int. Symp. Circuits Syst.* (Espoo, Finland), pp. 1795–1798.

J.W. Bandler, S. Daijavad and Q.J. Zhang (1985a), "Novel approach to multiplexer simulation and sensitivity analysis", *Electronics Letters*, vol.21, pp. 408–409.

J.W. Bandler, S. Daijavad and Q.J. Zhang (1985b), "Computer aided design of branched cascaded networks", *Proc. IEEE Int. Symp. Circuits Syst.* (Kyoto, Japan), pp. 1579–1582.

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, pp. 93–102.

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-R.

J.W. Bandler and 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 and W. Kellermann and K. Madsen (1987), "A nonlinear ℓ_1 optimization algorithm for design, modeling and diagnosis of networks", *IEEE Trans. Circuits Syst.*, vol. CAS-34, pp. 174-181.

J.W. Bandler and P.C. Liu (1974a), "Automatic network design with optimal tolerances", *IEEE Trans. Circuits Syst.*, vol. CAS-21, pp. 219-222.

J.W. Bandler and P.C. Liu (1974b), "Some implications of biquadratic functions in the tolerance problem", *Proc. IEEE Int. Symp. Circuits Syst.* (San Francisco, CA), pp. 740-744.

J.W. Bandler and P.C. Liu (1975), "Some implications of biquadratic functions in the tolerance problem", *IEEE Trans. Circuits Syst.*, vol. CAS-22, pp. 385-390.

J.W. Bandler, P.C. Liu and J.H.K. Chen (1974), "Computer-aided tolerance optimization applied to microwave circuits", *IEEE Int. Microwave Symp. Dig.* (Atlanta, GA), pp. 275-277.

J.W. Bandler, P.C. Liu and H. Tromp (1975), "Practical design centering, tolerancing and tuning", *Proc. IEEE Int. Symp. Circuits Syst.* (Newton, MA), pp. 206-209.

J.W. Bandler, P.C. Liu and H. Tromp (1976), "Integrated approach to microwave design ", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-24, pp. 584-591.

J.W. Bandler and Q.J. Zhang (1987), "An automatic decomposition technique for device modeling and large circuit design", *IEEE Int. Microwave Symp. Dig.* (Las Vegas, NV), pp. 709-712.

J.W. Bandler, Q.J. Zhang and R.M. Biernacki (1988a), "A unified framework for harmonic balance simulation and sensitivity analysis", *IEEE Int. Microwave Symp. Dig.* (New York, NY), pp. 1041-1044.

J.W. Bandler, Q.J. Zhang and R.M. Biernacki (1988b), "A unified theory for frequency-domain simulation and sensitivity analysis of linear and nonlinear circuits", *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 1661-1669.

J.W. Bandler, Q.J. Zhang and R.M. Biernacki (1989), "Practical, high-speed gradient computation for harmonic balance simulators", *IEEE Int. Microwave Symp. Dig.* (Long Beach, CA), pp. 363-366.

J.W. Bandler, Q.J. Zhang and Q. Cai (1990), "Nonlinear circuit optimization with dynamically integrated physical device models", *IEEE Int. Microwave Symp. Dig.* (Dallas, TX), pp. 567-570.

J.W. Bandler, Q.J. Zhang, J. Song and R.M. Biernacki (1989), "Yield optimization of nonlinear circuits with statistically characterized devices", *IEEE Int. Microwave Symp. Dig.* (Long Beach, CA), pp. 649–652.

J.W. Bandler, Q.J. Zhang, J. Song and R.M. Biernacki (1990), "FAST gradient based yield optimization of nonlinear circuits," *IEEE Trans. Microwave Theory Tech.*, vol. 38, pp. 1701–1710.

J.W. Bandler, Q.J. Zhang, S. Ye and S.H. Chen (1989), "Efficient large-signal FET parameter extraction using harmonics", *IEEE Trans. Microwave Theory Tech.*, vol. 37, pp. 2099–2108.

J.A. Barby, J. Vlach and K. Singhal (1988), "Polynomial splines for MOSFET model approximation," *IEEE Trans. Computer-Aided Design*, vol. 7, pp. 557–566.

R.M. Biernacki, J.W. Bandler, J. Song and Q.J. Zhang (1989), "Efficient quadratic approximation for statistical design", *IEEE Trans. Circuits Syst.*, vol. 36, pp. 1449–1454.

R.M. Biernacki and M.A. Styblinski (1986), "Statistical circuit design with a dynamic constraint approximation scheme", *Proc. IEEE Int. Symp. Circuits Syst.* (San Jose, CA), pp. 976–979.

G. Box, W. Hunter and J. Hunter (1978), *Statistics for Experimenters*. New York: Wiley, 1978.

R.K. Brayton, S.W. Director and G.D. Hachtel (1980), "Yield maximization and worst case design with arbitrary statistical distributions", *IEEE Trans. Circuits Syst.*, vol. CAS-27, pp. 756–764.

J.L. Burns, A.R. Newton and D.O. Pederson (1983), "Active device table look-up models for circuit simulation", *Proc. IEEE Int. Symp. Circuits Syst.* (Newport Beach, CA), pp. 250–252.

E.M. Butler (1971a), "Realistic design using large-change sensitivities and performance contours", *IEEE Trans. Circuit Theory*, vol. CT-18, pp. 58–66.

E.M. Butler (1971b), "Large change sensitivities for statistical design", *BSTJ*, vol. 50, pp. 1209–1224. Also in *Computer-Aided Filter Design*, G. Szentirmai, Ed., New York: IEEE Press, 1973.

D.A. Calahan (1979), "Vector processors-models and applications", *IEEE Trans. Circuits Syst.*, vol. CAS-26, pp. 715–726.

D.A. Calahan (1980), "Multi-level vectorized sparse solution of LSI circuits", *Proc. ICCC'80* (Port Chester, NY), pp. 976–979.

G. Casinovi and A. Sangiovanni-Vincentelli (1991), "A macromodeling algorithm for

analog circuits", *IEEE Trans. Computer-Aided Design*, vol. 10, pp. 150-160.

A. Casotto, F. Romeo and A. Sangiovanni-Vincentelli (1987), "A parallel simulated annealing algorithm for the placement of macro-cell", *IEEE Trans. Computer-Aided Design*, vol. CAD-6, pp. 838-847.

M.H. Chen (1983), "A 12-channel contiguous-band multiplexer at KU-band", 1983 *IEEE Int. Microwave Symp. Dig.* (Boston, MA), pp. 77-79.

M.H. Chen (1985), "Current state-of-the-art technology on contiguous-band multiplexers", *Proc. IEEE Int. Symp. Circuits Syst.* (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.

L.O. Chua and P-M. Lin (1975), *Computer-Aided Analysis of Electronic Circuits*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

P. Cox, P. Yang, S.S. Mahant-Shetti and P. Chatterjee (1985), "Statistical modeling for efficient parametric yield estimation of MOS VLSI circuits", *IEEE J. Solid-State Circuits*, vol. SC-20, pp. 391-398.

W. R. Curtice (1987), "Nonlinear analysis of GaAs MESFET amplifiers, mixers, and distributed amplifiers using the harmonic balance technique", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-35, pp. 441-447.

W. R. Curtice (1989), "Intrinsic GaAs MESFET equivalent circuit models generated from two-dimensional simulations", *IEEE Trans. Computer-Aided Design*, vol. 8, pp. 395-402.

S. W. Director and G. D. Hachtel (1977), "The simplicial approximation approach to design centering", *IEEE Trans. Circuits Syst.*, vol. CAS-24, pp. 363-372.

T. Downs, A.S. Cook and G. Rogers (1984), "A partitioning approach to yield estimation for large circuits and systems", *IEEE Trans. Circuits Syst.*, vol. CAS-31, pp. 472-485.

R. Gilmore (1986), "Nonlinear circuit design using the modified harmonic balance algorithm", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-34, pp. 1294-1307.

R. Gilmore and M. Steer (1991), "Nonlinear circuit analysis using the method of harmonic balance — a review of the art. II. advanced concepts", *J. Microwave and Millimeter-Wave Computer-Aided Engineering*, vol. 1, pp. 159-180.

K.C. Gupta, R. Garg and R. Chadha (1981), *Computer-Aided Design of Microwave Circuits*. Norwood, MA: Artech House Inc..

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.

N. Herr and J.J. Barnes (1986), "Statistical circuit simulation modeling of CMOS VLSI", *IEEE Trans. Computer-Aided Design*, vol. CAD-5, pp. 15–22.

D.E. Hocevar, M.R. Lightner and T.N. Trick (1983), "Monte Carlo based yield maximization with quadratic model", *Proc. IEEE Int. Symp. Circuits Syst.* (Newport Beach, CA), pp. 550–553.

N.K. Jain, D. Agnew and M.S. Nakhla (1983), "Two-dimensional table look-up MOSFET model", *Proc. Int. Conf. Computer-Aided Design* (Santa Clara, CA), pp. 201–203.

D. Jepsen and C. Gelatt (1983), "Macro placement by Monte Carlo annealing", *Proc. Int. Conf. Computer-Aided Design* (Santa Clara, CA), pp. 495–598.

B.J. Karafin (1971), "The optimal assignment of component tolerances for electrical networks", *BSTJ*, vol. 50, pp. 1225–1242. Also in *Computer-Aided Filter Design*, G. Szentirmai, Ed., New York: IEEE Press, 1973.

B.J. Karafin (1974), "The general component tolerance assignment problem in electrical networks", Ph.D. Thesis, Univ. of Pennsylvania, Philadelphia, PA.

M.A. Khatibzadeh and R.J. Trew (1988), "A large-signal, analytic model for the GaAs MESFET", *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 231–238.

S. Kirkpatrick, C.D. Gelatt and M.P. Vecchi (1983), "Optimization by simulated annealing," *Science*, vol. 220, pp. 671–680

K.S. Kundert and A. L. Sangiovanni-Vincentelli (1986), "Simulation of nonlinear circuits in the frequency domain", *IEEE Trans. Computer-Aided Design*, vol. CAD-5, pp. 521–535.

P.H. Ladbrooke (1989), *MMIC Design: GaAs FETs and HEMTs*. Norwood, MA: Artech House, 1989.

M.R. Lightner and S.W. Director (1981), "Multiple criterion optimization with yield maximization", *IEEE Trans. Circuits Syst.*, vol. CAS-28, pp. 781–791.

A. Materka and T. Kacprzak (1985), "Computer calculation of large-signal GaAs FET amplifier characteristics", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-33, pp. 129–135.

J. McDonald, R. Maini, L. Spangler and H. Weed (1989), "Response surface methodology: a modeling tool for integrated circuit designers", *IEEE J. Solid-State*

Circuits, vol. 24, pp. 469–473.

Microwave Harmonica User's Manual (1987), Compact Software Inc., Paterson, NJ 07504

L.J. Opalski and M.A. Styblinski (1986), "Generalization of yield optimization problem: maximum income approach", *IEEE Trans. Computer-Aided Design*, vol. CAD–5, pp. 346–360.

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 Syst.*, vol. CAS–26, pp. 795–813.

M.J.D. Powell (1978), "The convergence of variable metric method for nonlinearly constrained optimization calculations", in *Nonlinear Programming 3*, O.L. Mangasarian, R.R. Meyer and S.M. Robinson, Eds. New York. Academic Press, pp. 27–63.

J. Purviance, D. Criss and D. Monteith (1988), "FET model statistics and their effects on design centering and yield prediction for microwave amplifiers", *IEEE Int. Microwave Symp. Dig.*, (New York, NY), pp. 315–318.

J. Purviance and M. Meehan (1988), "A sensitivity figure for yield improvement", *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 413–417.

J. Purviance and M. Meehan (1991), "CAD for statistical analysis and design of microwave circuits", *Int. J. Microwave and Millimeter-Wave Computer-Aided Engineering*, vol. 1, pp. 59–76.

V. Rizzoli, C. Cecchetti and A. Neri (1986), "Supercomputer–aided generalized mixer analysis and optimization", *Proc. 16th European Microwave Conf.* (Dublin, Ireland), pp. 692–687.

V. Rizzoli, M. Ferlito and A. Neri (1986), "Vectorized program architectures for supercomputer–aided circuit design", *IEEE Trans. Microwave Theory Tech.*, vol. MTT–34, pp. 135–141.

V. Rizzoli, A. Lipparini, A. Costanzo, C. Cecchetti, A. Neri, F. Mastri, F. Sgallari, and V. Frontini (1991), "Modern perspectives in supercomputer–aided microwave circuit design", *Int. J. Microwave and Millimeter-Wave Computer-Aided Engineering*, vol. 1, pp. 201–224.

V. Rizzoli, A. Lipparini and E. Marazzi (1983), "A general–purpose program for nonlinear microwave circuit design", *IEEE Trans. Microwave Theory Tech.*, vol. MTT–31, pp. 762–770.

V. Rizzoli and A. Neri (1988), "State of the art and present trends in nonlinear microwave CAD techniques", *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 343–365.

F. Severson and S. Simpkins (1987), "Hadamard analysis—an effective, systematic tool for worst case circuit analysis", *Proc. IEEE CICC.* (Portland, OR), pp. 114–118.

K. Singhal and J.F. Pinel (1981), "Statistical design centering and tolerancing using parametric sampling", *IEEE Trans. Circuits Syst.*, vol. CAS-28, pp. 692–701.

C.M. Snowden and D. Loret (1987), "Two-dimensional hot-electron models for short-gate-length GaAs MESFET's", *IEEE Trans. Electron Devices*, vol. ED-34, pp. 212–223.

R. S. Soin and R. Spence (1978), "Statistical design centering for electrical circuits", *Electronics Letter*, vol. 14, pp. 772–774.

R. S. Soin and R. Spence (1980), "Statistical exploration approach to design centering", *Proc. Inst. Elec. Eng.*, vol. 127, pt. G, pp. 260–269.

R. S. Soin and R. Spence (1988), *Tolerance Design of Electronic Circuits*. New York, NY: Addison-Wesley, 1988.

M.L. Stein (1986), "An efficient method of sampling for statistical circuit design", *IEEE Trans. Computer-Aided Design*, vol. CAD-5, pp. 23–25.

A.J. Strojwas (1987), *Statistical Design of Integrated Circuits*. New York, NY: IEEE Press, 1987.

A.J. Strojwas (1989), "Design for manufacturability and yield", *26th ACM/IEEE Design Automation Conf.* (Las Vegas, NV), pp. 454–459.

A.J. Strojwas and A.L. Sangiovanni-Vincentelli Eds. (1986), *IEEE Trans. Computer-Aided Design*, Special Issue on Statistical Design of VLSI Circuits, vol. CAD-5, pp. 5–169.

M.A. Styblinski (1986), "Problems of yield gradient estimation for truncated probability density functions", *IEEE Trans. Computer-Aided Design*, vol. CAD-5, pp. 30–38.

M.A. Styblinski and A. Ruszczynski (1983), "Stochastic approximation approach to statistical circuit design", *Electronins Letter*, vol. 19, pp. 300–302.

Super-Compact User's Manual (1987), Compact Software Inc., Paterson, NJ 07504.

K.S. Tahim and R. Spence (1979),"A radial exploration algorithm for the statistical analysis of linear circuits", *IEEE Trans. Circuits Syst.*, vol. CAS-27, pp. 421–425.

A.R. Thorbjornsen and S.W. Director (1973), "Computer-aided tolerance assignment for linear circuits with correlated elements", *IEEE Trans. Circuit Theory*, vol. CT-20, pp. 518-523.

M. Vai, S. Prasad, N.C. Li and F. Kai (1989), "Modeling of microwave semiconductor devices using simulated annealing optimization", *IEEE Trans. Electron Devices*, vol. 36, pp. 761-762.

M. Vai, S. Prasad and B. Meskoob (1990), "Computer-aided design of monolithic distributed amplifiers with yield optimization", *IEEE Int. Microwave Symp. Dig.*, (Dallas, TX), pp. 347-350.

G.D. Vendelin, A.M. Pavio and U.L. Rohde (1990), *Microwave Circuit Design Using Linear and Nonlinear Techniques*. New York: Wiley, 1990.

J. Vlach and K. Singhal (1983), *Computer Methods for Circuit Analysis and Design*. New York: Van Nostrand Reinhild, 1983.

A. Vladimirescu and D.O. Pederson (1982), "Circuit simulation on vector processors", *Proc. ICCC'82* (New York, NY), pp. 172-175.

E. Wehrhahn and R. Spence (1984), "The performance of some design centering methods", *Proc. IEEE Int. Symp. Circuits Syst.* (Montreal, Canada), pp. 1424-1438.

F. Yamamoto and S. Takahashi (1985), "Vectorized LU decomposition algorithms for large-scale circuit simulation", *IEEE Trans. Computer-Aided Design*, vol. CAD-4, pp. 232-238.

A. Yoshii, M. Tomizawa and K. Yokoyama (1983), "Accurate modeling for submicrometer-gate Si and GaAs MESFET's using two-dimensional particle simulation", *IEEE Trans. Electron Devices*, vol. ED-30, pp. 1376-1380.

T. Yu, S.M. Kang, I.N. Hajj and T.N. Trick (1987), "Statistical performance modeling and parametric yield estimation of MOS VLSI", *IEEE Trans. Computer-Aided Design*, vol. CAD-6, pp. 1013-1022.

AUTHOR INDEX

SUBJECT INDEX