FAST Gradient Based Yield Optimization of Nonlinear Circuits

JOHN W. BANDLER, FELLOW, IEEE, QI-JUN ZHANG, MEMBER, IEEE, JIAN SONG, AND RADOSLAW M. BIERNACKI, SENIOR MEMBER, IEEE

***Abstract*—This paper meets the challenge of yield optimization of nonlinear microwave circuits operating in the steady state under large-signal periodic excitations. Yield-driven design is formulated as a one-sided ℓ_1 optimization problem. We introduce two novel, high-speed methods of gradient calculation, the integrated gradient approximation technique (IGAT) and the feasible adjoint sensitivity technique (FAST). IGAT utilizes the Broyden formula with special iterations of Powell to update the approximate gradients. FAST combines the efficiency and accuracy of the adjoint sensitivity technique with the simplicity of the perturbation technique. IGAT and FAST are compared with the simple perturbation approximate sensitivity technique (PAST) on the one extreme and the theoretical exact adjoint sensitivity technique (EAST) on the other. 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. A FET frequency doubler example treats statistics of both linear elements and nonlinear device parameters. This design has six optimizable variables including input power and bias conditions, and 34 statistical parameters. Using either IGAT or FAST, yield is driven from 40% to 70%. FAST exhibits superior efficiency.**

I. Introduction

STATISTICAL circuit design has been recognized as an indispensable tool for modern CAD of integrated circuits [1]–[3]. A number of algorithms for yield optimization have been developed within the past 15 years, e.g., Director and Hachtel [4] (simplicial approximation), Soin and Spence [5] (the center of gravity method), Bandler and Abdel-Malek [6], [7] (updated approximations and cuts), Styblinski and Ruszczynski [8] (stochastic approximation), Polak and Sangiovanni-Vincentelli [9] (outer approximation), Singhal and Pinel [10] (parametric sampling), Bandler and Chen [3] (generalized ℓ_p centering), and Biernacki *et al.* [11] (efficient quadratic approximation). This paper deals with yield optimization of nonlinear microwave circuits within the harmonic balance (HB) simulation environment.

Manuscript received March 6, 1990; revised July 12, 1990. This work was supported in part by Optimization Systems Associates Inc. and in part by the Natural Sciences and Engineering Research Council of Canada under Grants OGP0007239, OGP0042444, and STR0040923.

J. W. Bandler, Q. Zhang, and R. M. Biernacki are with Optimization Systems Associates Inc., P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7, and with the Simulation Optimization Systems Research Laboratory and the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada L8S 4L7.

J. Song is with the Simulation Optimization Systems Research Laboratory and the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada L8S 4L7.

IEEE Log Number 9038712.

Statistical design of practical nonlinear microwave circuits is a challenge. One serious inherent difficulty is the potentially prohibitively high computational cost: many circuits have to be simulated repeatedly and each circuit simulation involves CPU-intensive iterations to solve the HB 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 utmost importance.

The conventional perturbation approximate sensitivity technique (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 *et al.* [12] used this method in their single-loop approach for nominal circuit design. In yield optimization, however, PAST becomes extremely inefficient because of the large number of circuit outcomes to be dealt with.

The exact adjoint sensitivity technique (EAST) has been recently developed by Bandler, Zhang, and Biernacki [13], [14] for the HB 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.

In this paper, we formulate the yield-driven design of nonlinear circuits as a one-sided ℓ_1 optimization problem [15] allowing the powerful, robust one-sided ℓ_1 algorithm proposed by Bandler *et al.* [16] to be employed. We introduce two powerful approaches to gradient calculation. One is the integrated gradient approximation technique (IGAT), presented by Bandler *et al.* [17], which is adapted here to the needs of yield optimization. The other is the feasible adjoint sensitivity technique (FAST), first reported by Bandler *et al.* [18]. Motivated by the potential impact of the adjoint sensitivity approach on general-purpose CAD programs, we have studied its implementational aspects. 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.

IGAT and FAST are applied to yield optimization of a microwave frequency doubler. In this example, normal and uniform distributions describing large-signal FET model parameters and passive elements are fully accommodated. The performance yield is increased from 40% to 70%.

In Section II we formulate the yield optimization problem for nonlinear circuits as a one-sided ℓ_1 optimization problem. Sections III and IV are devoted to IGAT and FAST, respectively. Comparisons between the various approaches are made in Section V. Section VI presents the details of the FET frequency doubler example.

II. The Yield Problem for Nonlinear Circuits

A. 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. Therefore, design specifications can be imposed at dc and several harmonics. The jth specification can be denoted by

$$S_{uj}(\boldsymbol{h}) \tag{1a}$$

if it is an upper specification, or

$$S_{lj}(\boldsymbol{h}) \tag{1b}$$

in the case of lower specifications, where

$$\boldsymbol{h} = [0 \quad 1 \quad 2 \quad \cdots \quad H]^T \tag{2}$$

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

Manufactured outcomes are spread over a region which can be described by the nominal design, $\boldsymbol{\phi}^0$, along with a statistical distribution of parameters. Parameters in $\boldsymbol{\phi}^0$ can be lumped element values, device model parameters, dimensions of a physical realization, etc. For statistical design, many circuits are needed to represent the distribution of manufactured outcomes. Such circuit outcomes, denoted by $\boldsymbol{\phi}^i$, can be written as

$$\boldsymbol{\phi}^i = \boldsymbol{\phi}^0 + \Delta\boldsymbol{\phi}^i, \qquad i = 1, 2, \cdots, N \tag{3}$$

where $\Delta\boldsymbol{\phi}^i$ is the deviation of the ith outcome from the nominal circuit and N is the number of outcomes considered. In yield optimization, each $\boldsymbol{\phi}^i$ is determined by statistically perturbing $\boldsymbol{\phi}^0$ according to a known or assumed statistical distribution of the manufactured outcomes.

The response of each outcome, denoted by

$$R_j(\boldsymbol{\phi}^i, \boldsymbol{h}) \tag{4}$$

is calculated after solving the HB equations [19]

$$\bar{\boldsymbol{F}}(\boldsymbol{\phi}^i, \bar{\boldsymbol{V}}) = \boldsymbol{0} \tag{5}$$

where $\bar{\boldsymbol{V}}$ comprises the split real and imaginary parts of the state variables in the HB equation. The corresponding error function is defined as

$$R_j(\boldsymbol{\phi}^i, \boldsymbol{h}) - S_{uj}(\boldsymbol{h}) \tag{6a}$$

or as

$$S_{lj}(\boldsymbol{h}) - R_j(\boldsymbol{\phi}^i, \boldsymbol{h}). \tag{6b}$$

We assemble all errors for the outcome $\boldsymbol{\phi}^i$ into one vector $\boldsymbol{e}^i$. If all entries of this vector are nonpositive, the outcome $\boldsymbol{\phi}^i$ represents an acceptable circuit.

For N statistical outcomes generated, the production yield can be estimated by

$$Y \approx \frac{\text{number of acceptable circuits}}{N}. \tag{7}$$

B. Formulation of One-Sided ℓ_1 Objective Function

The problem of yield optimization can be properly converted to a mathematical programming problem so that modern mathematical optimization techniques can be applied. In the following the design variables are nominal values $\boldsymbol{\phi}^0$. Although only the outcomes $\boldsymbol{\phi}^i$ appear in the error functions, they depend on $\boldsymbol{\phi}^0$ because the $\boldsymbol{\phi}^i$ are related to $\boldsymbol{\phi}^0$.

After the error vector $\boldsymbol{e}^i$ for the outcome $\boldsymbol{\phi}^i$ has been assembled as

$$\boldsymbol{e}^i = [e_1(\boldsymbol{\phi}^i) \quad e_2(\boldsymbol{\phi}^i) \cdots e_M(\boldsymbol{\phi}^i)]^T \tag{8}$$

where M is the total number of errors considered, the formulation of the objective function for optimization can follow the procedure described in [3]. First, we create the generalized ℓ_p function v^i from $\boldsymbol{e}^i$ [3]:

$$v^i = \begin{cases} \left[\sum_{j \in J(\boldsymbol{\phi}^i)} \left(e_j(\boldsymbol{\phi}^i)\right)^p \right]^{1/p} & \text{if } J(\boldsymbol{\phi}^i) \neq \varnothing \quad \text{(9a)} \\ -\left[\sum_{j=1}^{M} \left(-e_j(\boldsymbol{\phi}^i)\right)^{-p} \right]^{-1/p} & \text{if } J(\boldsymbol{\phi}^i) = \varnothing \quad \text{(9b)} \end{cases}$$

where

$$J(\boldsymbol{\phi}^i) = \left\{ j \,\middle|\, e_j(\boldsymbol{\phi}^i) \geq 0 \right\}. \tag{10}$$

Then we define the one-sided ℓ_1 objective function for yield optimization [3] as

$$U(\boldsymbol{\phi}^0) = \sum_{i \in I} \alpha_i v^i \tag{11}$$

where

$$I = \{ i \mid v^i > 0 \} \tag{12}$$

and α_i are positive multipliers. If the α_i were chosen as [3]

$$\alpha_i = \frac{1}{|v^i|} \tag{13}$$

then the function $U(\boldsymbol{\phi}^0)$ would become the exact number of unacceptable circuits and the yield would be

$$Y(\phi^0) = 1 - \frac{U(\phi^0)}{N}. \tag{14}$$

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,

$$\underset{\phi^0}{\text{minimize}}\, U(\phi^0). \tag{15}$$

We use (13) 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.

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

$$U(\phi^0) = \sum_{i \in I} \sum_{j \in J(\phi^i)} \alpha_i e_j(\phi^i) \tag{16}$$

where α_i, I, and $J(\phi^i)$ are defined as before. In (16), error functions for optimization are $e_j(\phi^i)$. In (11), error functions for optimization are v^i. The functions $e_j(\phi^i)$ are differentiable, but the functions v^i of (9) may not be. Therefore, we use (16) in our yield optimization.

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.

C. The One-Sided ℓ_1 Optimization Algorithm

We use a highly efficient one-sided ℓ_1 optimization algorithm [16] to solve (15). 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 done to ensure global convergence of the combined algorithm.

To summarize the discussion in this section, all steps involved in our yield optimization are shown in Fig. 1.

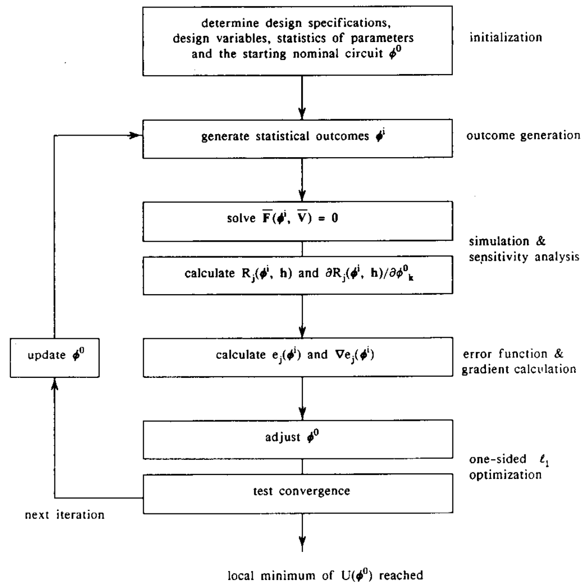


Fig. 1. Flowchart for yield optimization.

III. Integrated Gradient Approximation Technique

After a review of PAST, IGAT is discussed for the case of a single function. The application of IGAT in yield optimization follows.

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

A. Approximating Derivatives by PAST

The first-order derivative of $f(\phi)$ with respect to the kth variable can be estimated by

$$\frac{\partial f(\phi)}{\partial \phi_k} \approx \frac{f(\phi + \Delta\phi_k \boldsymbol{u}_k) - f(\phi)}{\Delta\phi_k} \tag{17}$$

where $\phi + \Delta\phi_k \boldsymbol{u}_k$ denotes the perturbation of the kth variable, $\Delta\phi_k$ is the perturbation length and $\boldsymbol{u}_k$ 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.

B. IGAT for General Functions [17]

To start the process, PAST is used as in (17) to calculate the approximate gradient.

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

$$\nabla f(\phi_{\text{new}}) = \nabla f(\phi_{\text{old}}) + \frac{f(\phi_{\text{new}}) - f(\phi_{\text{old}}) - (\nabla f(\phi_{\text{old}}))^T \Delta\phi}{\Delta\phi^T \Delta\phi} \Delta\phi \tag{18}$$

where ϕ_{old} and ϕ_{new} are two different points and $\Delta\phi = \phi_{\text{new}} - \phi_{\text{old}}$. If ϕ_{old} and ϕ_{new} are iterates of optimization, $f(\phi_{\text{old}})$ and $f(\phi_{\text{new}})$ 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 (17) to maintain the accuracy of the approximate gradients at a desirable level.

C. Application of IGAT in Yield Optimization

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. We still use f to denote a particular response function whose gradient is to be approximated.

Because the nominal values, $\boldsymbol{\phi}^0$, are design variables, all perturbations are made to $\boldsymbol{\phi}^0$ in the initialization and reinitialization steps using PAST (17). When $\boldsymbol{\phi}^0$ is perturbed to $\boldsymbol{\phi}^0 + \Delta\phi_k^0 \boldsymbol{u}_k$, denoted by $\boldsymbol{\phi}_{k,\text{pert}}^0$ for short, outcomes should be regenerated from $\boldsymbol{\phi}_{k,\text{pert}}^0$ in order to get perturbed circuit responses. These outcomes are denoted by $\boldsymbol{\phi}_{k,\text{pert}}^i$. Then from (17), the approximate derivative of the response $f(\boldsymbol{\phi}^i)$ is defined as

$$\frac{\partial f(\boldsymbol{\phi}^i)}{\partial \phi_k^0} \approx \frac{f(\boldsymbol{\phi}_{k,\text{pert}}^i) - f(\boldsymbol{\phi}^i)}{\Delta\phi_k^0}. \tag{19}$$

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

$$\nabla f(\boldsymbol{\phi}_{\text{new}}^i) = \nabla f(\boldsymbol{\phi}_{\text{old}}^i) + \frac{f(\boldsymbol{\phi}_{\text{new}}^i) - f(\boldsymbol{\phi}_{\text{old}}^i) - (\nabla f(\boldsymbol{\phi}_{\text{old}}^i))^T \Delta\boldsymbol{\phi}^0}{(\Delta\boldsymbol{\phi}^0)^T \Delta\boldsymbol{\phi}^0} \Delta\boldsymbol{\phi}^0. \tag{20}$$

IV. Feasible Adjoint Sensitivity Technique

In the HB simulation environment, the sensitivity of a response with respect to one variable, ϕ_k,

$$\frac{\partial R_j(\boldsymbol{\phi}, \boldsymbol{h})}{\partial \phi_k} \tag{21}$$

should be computed subject to the constraints of the HB equation, e.g., [19]

$$\bar{\boldsymbol{F}}(\boldsymbol{\phi}, \bar{\boldsymbol{V}}) = \boldsymbol{0}. \tag{22}$$

Bandler *et al.* [13], [14] proposed EAST, and the sensitivity expressions for various elements were derived and listed [14].

Here, we propose FAST, which is also based on adjoint sensitivity principles. Suppose that the circuit is divided into linear and nonlinear subnetworks and that the response of interest is the voltage at the output port. Normally, the response is taken from the linear subnetwork.

The response can be calculated by

$$\bar{V}_{\text{out}} = [\boldsymbol{a}^T \quad \boldsymbol{b}^T] \begin{bmatrix} \bar{\boldsymbol{V}} \\ \bar{\boldsymbol{V}}_s \end{bmatrix} = \boldsymbol{c}^T \begin{bmatrix} \bar{\boldsymbol{V}} \\ \bar{\boldsymbol{V}}_s \end{bmatrix} \tag{23}$$

where $\bar{\boldsymbol{V}}_s$ denotes the split real and imaginary parts in the spectra of excitation voltages, $\bar{\boldsymbol{V}}$ denotes the solution to the HB equations (22), and

$$\boldsymbol{c} = [\boldsymbol{a}^T \quad \boldsymbol{b}^T]^T \tag{24}$$

is a linear transfer vector linking the output voltage with $\bar{\boldsymbol{V}}_s$ and $\bar{\boldsymbol{V}}$. $\bar{\boldsymbol{V}}$ and $\boldsymbol{c}$ are functions of $\boldsymbol{\phi}$. $\bar{\boldsymbol{V}}_s$ can also be a function of $\boldsymbol{\phi}$ if we want to change $\bar{\boldsymbol{V}}_s$ to improve the circuit performance.

A. FAST for the Nominal Circuit Case

To make the derivation procedure concise, we concentrate on a single circuit design with variables $\boldsymbol{\phi}$. From (23) the approximate derivative of $\bar{V}_{\text{out}}$ w.r.t. ϕ_k can be calculated as

$$\frac{\Delta \bar{V}_{\text{out}}}{\Delta \phi_k} \approx \frac{\Delta \boldsymbol{c}^T}{\Delta \phi_k} \begin{bmatrix} \bar{\boldsymbol{V}} \\ \bar{\boldsymbol{V}}_s \end{bmatrix} + \boldsymbol{a}^T \frac{\Delta \bar{\boldsymbol{V}}}{\Delta \phi_k} + \boldsymbol{b}^T \frac{\Delta \bar{\boldsymbol{V}}_s}{\Delta \phi_k} \tag{25}$$

by perturbing $\boldsymbol{\phi}$ to $\boldsymbol{\phi} + \Delta\phi_k \boldsymbol{u}_k$. Let $\hat{\bar{\boldsymbol{V}}}$ be the adjoint voltages obtained by solving the set of linear equations

$$\boldsymbol{J}^T \hat{\bar{\boldsymbol{V}}} = \boldsymbol{a} \tag{26}$$

where $\boldsymbol{J}$ is the Jacobian of $\bar{\boldsymbol{F}}$ w.r.t. $\bar{\boldsymbol{V}}$ at the HB solution. We can express

$$\Delta \bar{V}_{\text{out}} \approx [\boldsymbol{c}^T(\boldsymbol{\phi} + \Delta\phi_k \boldsymbol{u}_k) - \boldsymbol{c}^T] \begin{bmatrix} \bar{\boldsymbol{V}} \\ \bar{\boldsymbol{V}}_s \end{bmatrix} + \boldsymbol{b}^T[\bar{\boldsymbol{V}}_s(\boldsymbol{\phi} + \Delta\phi_k \boldsymbol{u}_k) - \bar{\boldsymbol{V}}_s] - \hat{\bar{\boldsymbol{V}}}^T \Delta \bar{\boldsymbol{F}}. \tag{27}$$

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

$$\Delta \bar{\boldsymbol{F}} \approx \bar{\boldsymbol{F}}(\boldsymbol{\phi} + \Delta\phi_k \boldsymbol{u}_k, \bar{\boldsymbol{V}}) \tag{28}$$

for a small $\Delta\phi_k$.

Considering the different elements, (27) can be further expressed as

$$\Delta \bar{V}_{\text{out}} \approx \begin{cases} [\boldsymbol{c}^T(\boldsymbol{\phi} + \Delta\phi_k \boldsymbol{u}_k) - \boldsymbol{c}^T] \begin{bmatrix} \bar{\boldsymbol{V}} \\ \bar{\boldsymbol{V}}_s \end{bmatrix} - \hat{\bar{\boldsymbol{V}}}^T \bar{\boldsymbol{F}}(\boldsymbol{\phi} + \Delta\phi_k \boldsymbol{u}_k, \bar{\boldsymbol{V}}), & \phi_k \in \text{linear subnetwork} \\ \boldsymbol{b}^T[\bar{\boldsymbol{V}}_s(\boldsymbol{\phi} + \Delta\phi_k \boldsymbol{u}_k) - \bar{\boldsymbol{V}}_s] - \hat{\bar{\boldsymbol{V}}}^T \bar{\boldsymbol{F}}(\boldsymbol{\phi} + \Delta\phi_k \boldsymbol{u}_k, \bar{\boldsymbol{V}}), & \phi_k \in \text{sources} \\ -\hat{\bar{\boldsymbol{V}}}^T \bar{\boldsymbol{F}}(\boldsymbol{\phi} + \Delta\phi_k \boldsymbol{u}_k, \bar{\boldsymbol{V}}), & \phi_k \in \text{nonlinear subnetwork}. \end{cases} \tag{29}$$

This formula is much easier to implement than the corresponding formula for EAST [13], [14]. The function $\bar{F}(\phi + \Delta\phi_k u_k, \bar{V})$ is evaluated by perturbation. The effort for solving the linear equations (26) is small since the LU factors of the Jacobian matrix are already available from the final HB iteration. The terms $\bar{V}$ and $\bar{V}_s$ are also available from the HB 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 $\bar{V}_s(\phi + \Delta\phi_k u_k)$ can be effortlessly obtained. It is clear that the calculation of all the terms in (27) or (29) can be readily implemented.

Finally, the approximate sensitivity of output voltage $\bar{V}_{\text{out}}$ w.r.t. ϕ_k can be computed as

$$\frac{\partial \bar{V}_{\text{out}}}{\partial \phi_k} \approx \frac{\Delta \bar{V}_{\text{out}}}{\Delta \phi_k}. \tag{30}$$

B. FAST for Yield Optimization

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

$$\begin{aligned} \Delta\bar{V}_{\text{out}}(\phi^i) \approx & \left[c^T(\phi^i_{k,\text{pert}}) - c^T(\phi^i)\right]\begin{bmatrix} \bar{V} \\ \bar{V}_s(\phi^i)\end{bmatrix} \\ & + b^T(\phi^i)\left[\bar{V}_s(\phi^i_{k,\text{pert}}) - \bar{V}_s(\phi^i)\right] \\ & - \hat{\bar{V}}^T \Delta\bar{F} \end{aligned} \tag{31}$$

where

$$\Delta\bar{F} \approx \bar{F}(\phi^i_{k,\text{pert}}, \bar{V}). \tag{32}$$

$\bar{V}$ is the solution to the HB equation (5), and $\hat{\bar{V}}$ is the solution to (26).

V. Comparisons of Various Approaches

A. 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, the cost may be prohibitive. Suppose there are ten 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., ten times. The best possible situation for this approach is that all ten simulations use the same Jacobian and all converge in one iteration. This applies to nominal circuit design. For yield optimization, a large number of statistically generated circuit 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.

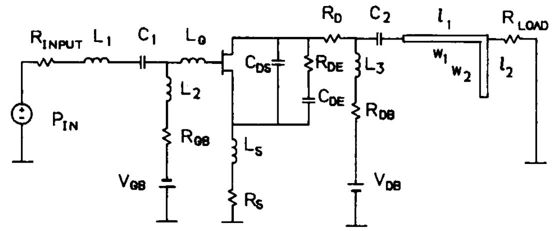


Fig. 2. Circuit diagram of the FET microwave frequency doubler. The nominal values for nonoptimizable variables are: $L_2 = 15$ nH, $L_3 = 15$ nH, $C_1 = 20$ pF, $C_2 = 20$ pF, $w_1 = 0.1 \times 10^{-3}$ m, $w_2 = 0.635 \times 10^{-3}$ m, $R_{\text{LOAD}} = R_{\text{INPUT}} = 50\ \Omega$, and $R_{GB} = R_{DB} = 10\ \Omega$.

Both EAST and FAST require modification to the circuit simulator.

The generic exact adjoint sensitivity technique [13], [14] 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 ten design variables, instead of completely solving ten nonlinear circuits, we only evaluate ten residuals in the form of (28) and calculate the perturbed linear subnetwork. The solution of adjoint equation (26) 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.

B. Numerical Comparison of FAST, EAST, and PAST

We use a MESFET mixer [13], [14] to investigate the accuracy and actual time efficiency of FAST [18]. Sensitivities of the mixer conversion gain w.r.t. 26 variables were calculated by the FAST, EAST, and PAST approaches, respectively. The variables included all parameters in the linear as well as the nonlinear part, dc bias, LO power, and 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% 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 three times slower than EAST but 23 times faster than PAST for one complete sensitivity analysis of the mixer circuit.

VI. Yield Optimization of a Frequency Doubler

The FET frequency doubler shown in Fig. 2 is considered. It consists of a common-source FET with a lumped input matching network and a microstrip output matching

TABLE I
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_l	0.676×10^{-1}	0.65
$R_D(\Omega)$	2.153	3	K_G	1.1	0.65
L_S(nH)	0.07	5	τ(pS)	7.0	6
$R_S(\Omega)$	1.144	5	S_S	1.666×10^{-3}	0.65
$R_{DE}(\Omega)$	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	$R_{10}(\Omega)$	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 [20].

TABLE II
FET MODEL PARAMETER CORRELATIONS [22]

	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 to be consistent with the large-signal FET model.

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_{GB} and V_{DB} and the driving power level P_{IN} are also considered as optimization variables. The fundamental frequency is 5 GHz. Responses of interest are the conversion gain and spectral purity, which are 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 2.5 dB and 20 dB, respectively. They are both lower specifications.

Our large-signal FET statistical model includes an intrinsic large-signal FET model modified from the Materka and Kacprzak model [21], 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 I. The correlations between parameters are assumed according to the results published by Purviance *et al.* [22]. Certain modifications have been made to make the correlations for the large-signal FET model consistent with those for the small-signal FET model dealt with in [22]. The correlation coefficients are given in Table II. Uniform distributions with fixed tolerances of 3% are assumed for P_{IN}, V_{GB}, V_{DB}, L_1, l_1, and l_2. Finally, 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.

In our program, the formulation (16) is used. In more detail, the error functions resulting from the simulated conversion gain and spectral purity are calculated, and then these error functions with their multipliers defined in (13) are fed into the one-side ℓ_1 optimizer. IGAT and FAST are implemented to provide gradients. IGAT calculates approximate sensitivities of the conversion gain and spectral purity. The computer used is a 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 two designs using IGAT and FAST gradient calculation in the same environment. Computational details are given in Tables III and IV. Each 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.

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. For FAST, 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 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 ap-

[1]A test at McMaster University using LINPACK to solve 100 linear equations indicated a performance for this machine of 10 MFLOPS. A VAX 11/780 exhibited 0.14 MFLOPS in a similar test.

TABLE III
YIELD OPTIMIZATION OF THE FREQUENCY DOUBLER USING IGAT

Variable	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_{GB}(V)	-1.9	-1.70329	-1.93468	-1.92542
V_{DB}(V)	5.0	6.50000	6.50000	6.50000
L_1(nH)	5.0	5.29066	5.68905	5.63822
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 Function Evaluations			89	95
CPU Time (Multiflow Trace 14/300)			18.6min	19.1min

The yield is estimated from 500 outcomes.

TABLE IV
YIELD OPTIMIZATION OF THE FREQUENCY DOUBLER USING FAST

Variable	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_{GB}(V)	-1.9	-1.70329	-1.93930	-1.92927
V_{DB}(V)	5.0	6.50000	6.50000	6.5000
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 Function Evaluations and Sensitivity Analyses			19	29
CPU Time (Multiflow Trace 14/300)			7.9min	12.1min

The yield is estimated from 500 outcomes.

proach. Although IGAT is slower than FAST, it is very robust in terms of the final yield reached.

Figs. 3 and 4 show histograms of the conversion gain and the spectral purity, respectively. Five hundred outcomes are used to calculate both distributions. Fig. 3(a) is the conversion gain distribution before yield optimization. The histograms in parts (b) and (c) of Fig. 3 are based on the solutions using IGAT and FAST, respectively. Fig. 4(a) is the spectral purity distribution before yield optimization. The histograms in parts (b) and (c) of Fig. 4 are based on solutions using IGAT and FAST, respectively. The improvement in spectral purity is very clearly illustrated by the histograms in Fig. 4. Before yield optimization, the center of the distribution is close to the design specification of 20 dB, 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.

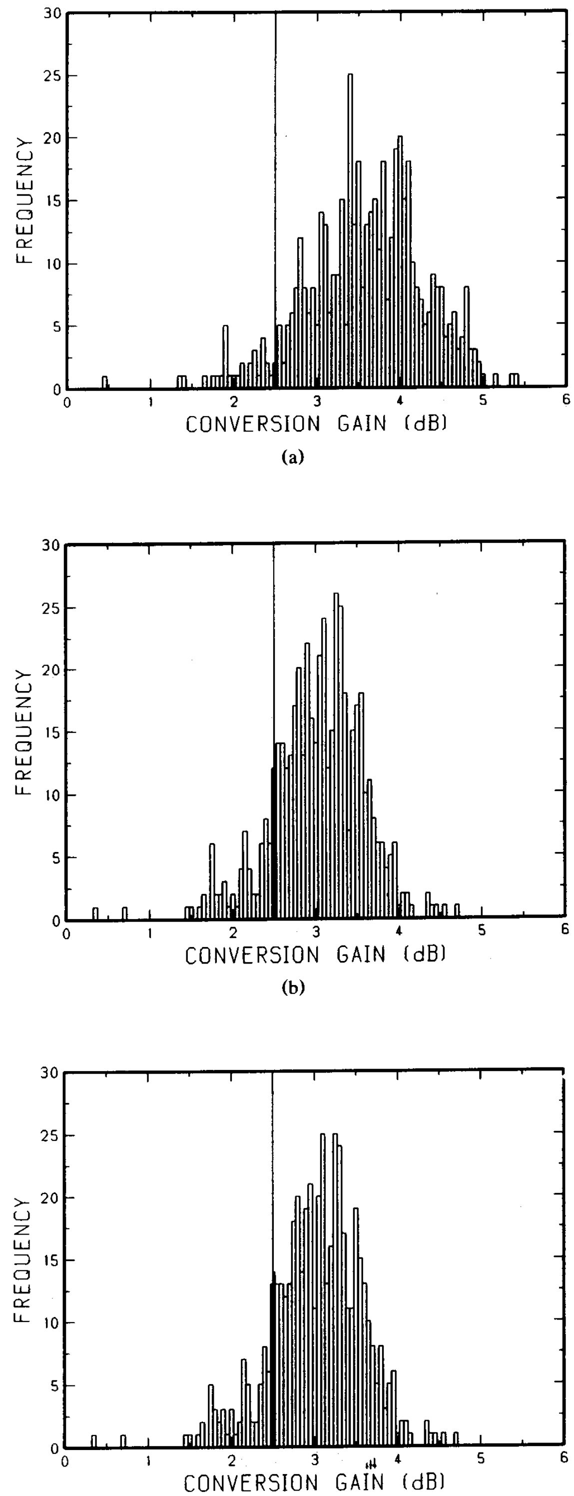


Fig. 3. Histogram of conversion gains of the frequency doubler based on 500 statistical outcomes. The specification is shown by a vertical line. (a) At the starting point. (b) At the solution of yield optimization using IGAT. (c) At the solution of yield optimization using FAST.

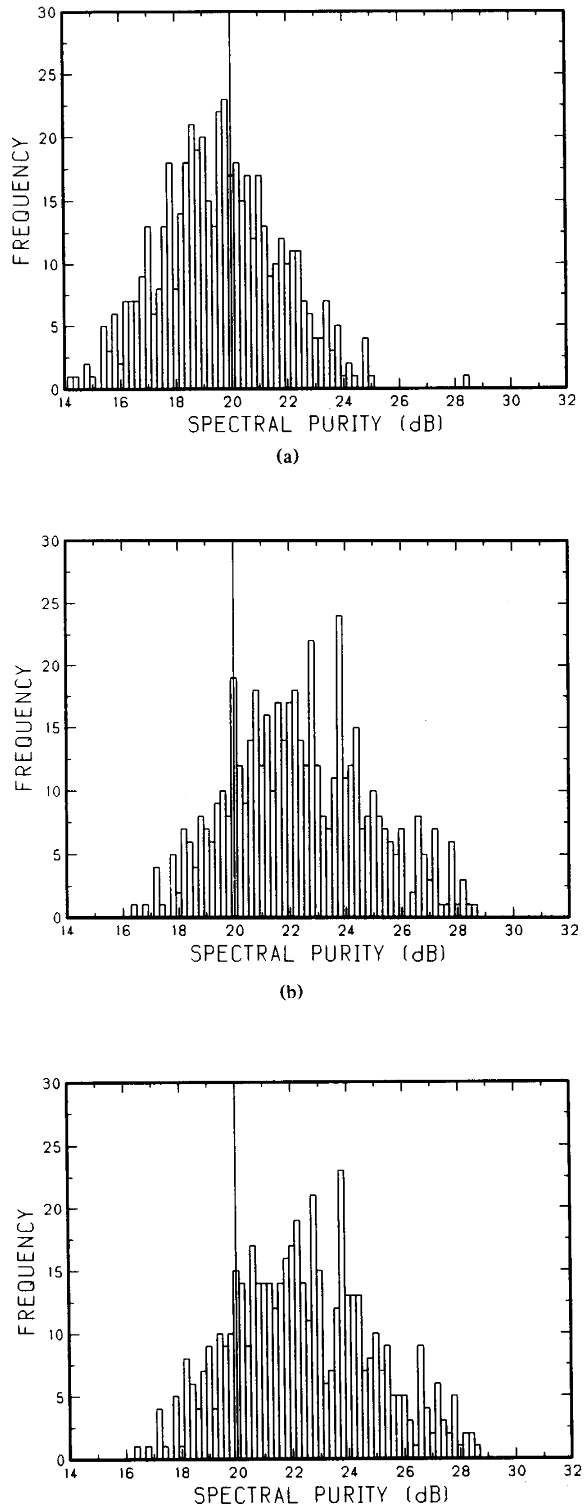


Fig. 4. Histogram of spectral purities of the frequency doubler based on 500 statistical outcomes. The specification is shown by a vertical line. (a) At the starting point. (b) At the solution of yield optimization using IGAT. (c) At the solution of yield optimization using FAST.

VII. Conclusions

This paper presents a comprehensive formulation for yield optimization of nonlinear circuits operating within the harmonic balance simulation environment. We have conducted a convincing demonstration of yield optimization of statistically characterized nonlinear microwave circuits using our two best approaches to gradient calculation, namely, IGAT and FAST. These two approaches are expedient tools for gradient calculation in the HB environment. The significant advantages of IGAT and FAST over PAST are their unmatched speeds, and over EAST are their 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 our two gradient calculation approaches. Large-signal FET parameter statistics are fully facilitated. The substantial computational advantages of IGAT and FAST have been observed. Our approaches provide powerful tools to meet the very pressing need for efficient microwave nonlinear circuit design. Our success should strongly motivate the development of statistical modeling of microwave devices for large-signal applications.

Acknowledgment

The authors wish to thank Dr. S. H. Chen of Optimization Systems Associates Inc., Dundas, Ontario, Canada, for providing some necessary software and suggestions.

References

[1] A. J. Strojwas, *Statistical Design of Integrated Circuits.* New York, NY: IEEE Press, 1987.
[2] E. Wehrhahn and R. Spence, "The performance of some design centering methods," in *Proc. IEEE Int. Symp. Circuits Syst.* (Montreal, Canada), 1984, pp. 1424–1438.
[3] J. W. Bandler and S. H. Chen, "Circuit optimization: The state of the art," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 424–443, 1988.
[4] S. W. Director and G. D. Hachtel, "The simplicial approximation approach to design centering," *IEEE Trans. Circuits Syst.*, vol. CAS-24, pp. 363–372, 1977.
[5] R. S. Soin and R. Spence, "Statistical exploration approach to design centering," *Proc. Inst. Elec. Eng.*, vol. 127, pt. G., pp. 260–269, 1980.
[6] J. W. Bandler and H. L. Abdel-Malek, "Optimal centering, tolerancing, and yield determination via updated approximations and cuts," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 853–871, 1978.
[7] H. L. Abdel-Malek and J. W. Bandler, "Yield optimization for arbitrary statistical distributions: Part I—Theory," *IEEE Trans. Circuits Syst.*, vol. CAS-27, pp. 245–253, 1980.
[8] M. A. Styblinski and A. Ruszczynski, "Stochastic approximation approach to statistical circuit design," *Electron. Lett.*, vol. 19, no. 8, pp. 300–302, 1980.
[9] E. Polak and A. L. Sangiovanni-Vincentelli, "Theoretical and computational aspects of the optimal design centering, tolerancing, and tuning problem," *IEEE Trans. Circuits Syst.*, vol. CAS-26, pp. 795–813, 1979.
[10] K. Singhal and J. F. Pinel, "Statistical design centering and tolerancing using parametric sampling," *IEEE Trans. Circuits Syst.*, vol. CAS-28, pp. 692–701, 1981.
[11] R. M. Biernacki, J. W. Bandler, J. Song, and Q. J. Zhang, "Efficient quadratic approximation for statistical design," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 1449–1454, 1989.

[12] V. Rizzoli, A. Lipparini, and E. Marazzi, "A general-purpose program for nonlinear microwave circuit design," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-31, pp. 762–770, 1983.
[13] J. W. Bandler, Q. J. Zhang, and R. M. Biernacki, "A unified framework for harmonic balance simulation and sensitivity analysis," in *IEEE MTT-S Int. Microwave Symp. Dig.* (New York, NY), 1988, pp. 1041–1044.
[14] J. W. Bandler, Q. J. Zhang, and R. M. Biernacki, "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, 1988.
[15] J. W. Bandler, Q. J. Zhang, J. Song, and R. M. Biernacki, "Yield optimization of nonlinear circuits with statistically characterized devices," in *IEEE MTT-S Int. Microwave Symp. Dig.* (Long Beach, CA), 1989, pp. 649–652.
[16] J. W. Bandler, S. H. Chen, and K. Madsen, "An algorithm for one-sided ℓ_1 optimization with application to circuit design centering," in *Proc. IEEE Int. Symp. Circuits Syst.* (Espoo, Finland), 1988, pp. 1795–1798.
[17] J. W. Bandler, S. H. Chen, S. Daijavad, and K. Madsen, "Efficient optimization with integrated gradient approximations," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 444–454, 1988.
[18] J. W. Bandler, Q. J. Zhang, and R. M. Biernacki, "Practical, high-speed gradient computation for harmonic balance simulators," in *IEEE MTT-S Int. Microwave Symp. Dig.* (Long Beach, CA), 1989, pp. 363–367.
[19] K. S. Kundert and A. L. Sangiovanni-Vincentelli, "Simulation of nonlinear circuits in the frequency domain," *IEEE Trans. Computer-Aided Design*, vol. CAD-5, pp. 521–535, 1986.
[20] J. W. Bandler, Q. J. Zhang, S. Ye, and S. H. Chen, "Efficient large-signal FET parameter extraction using harmonics," *IEEE Trans. Microwave Theory Tech.*, vol. 37, pp. 2099–2108, 1989.
[21] A. Materka and T. Kacprzak, "Computer calculation of large-signal GaAs FET amplifier characteristics," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-33, pp. 129–135, 1985.
[22] J. Purviance, D. Criss, and D. Monteith, "FET model statistics and their effects on design centering and yield prediction for microwave amplifiers," in *IEEE MTT-S Int. Microwave Symp. Dig.* (New York, NY), 1988, pp. 315–318.

John W. Bandler (S'66–M'66–SM'74–F'78) was born in Jerusalem on November 9, 1941. He studied at the Imperial College of Science and Technology, London, England, from 1960 to 1966. He received the B.Sc. (Eng.), Ph.D., and D.Sc. (Eng.) degrees from the University of London, London, England, in 1963, 1967, and 1976, respectively.

He joined Mullard Research Laboratories, Redhill, Surrey, England, in 1966. From 1967 to 1969 he was a Postdoctorate Fellow and Sessional Lecturer at the University of Manitoba, Winnipeg, Canada. He joined McMaster University, Hamilton, Canada, in 1969, where he is currently a Professor of Electrical and Computer Engineering. Dr. Bandler has served as Chairman of the Department of Electrical Engineering and Dean of the Faculty of Engineering. He currently directs research in the Simulation Optimization Systems Research Laboratory. He is President of Optimization Systems Associates Inc., Dundas, Ontario, Canada, which he established in 1983. OSA currently provides consulting services and software to numerous microwave clients in CAE of microwave integrated circuits. He is also President of Bandler Research Inc., Dundas, Ontario, Canada, which he founded in 1989.

Dr. Bandler contributed to *Modern Filter Theory and Design* (Wiley-Interscience, 1973) and to *Analog Methods for Computer-Aided Analysis and Diagnosis* (Marcel Dekker, 1988). He has more than 230 publications, four of which appear in *Computer-Aided Filter Design* (IEEE Press, 1973), one in *Microwave Integrated Circuits* (Artech House, 1975), one in *Low-Noise Microwave Transistors and Amplifiers* (IEEE Press, 1981), one in *Microwave Integrated Circuits* (2nd ed., Artech House, 1985), and one in *Statistical Design of Integrated Circuits* (IEEE Press, 1987). Dr. Bandler was an Associate Editor of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES (1969–1974). He was Guest Editor of a special issue of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES on computer-oriented microwave practices (March 1974). Dr. Bandler is on the editorial boards of the *International Journal of Numerical Modelling* and the *International Journal of Microwave and Millimeterwave Computer-Aided Engineering*. Dr. Bandler is a Fellow of the Royal Society of Canada and of the Institution of Electrical Engineers (Great Britain). He is a member of the Association of Professional Engineers of the province of Ontario (Canada) and a member of the Electromagnetics Academy.

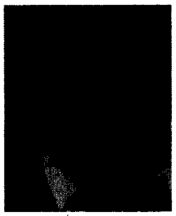
Qi-jun Zhang (S'84–M'87) was born in Xianyan, Shanxi, China, on October 8, 1959. He received the B.Eng. degree from the East China Engineering Institute, Nanjing, China, in 1982, and the Ph.D. degree from McMaster University, Hamilton, Canada, in 1987, both in electrical engineering.

He was a research assistant at the Institute of Systems Engineering, Tianjin University, Tianjin, China, from 1982 to 1983. He joined the Simulation Optimization Systems Research Laboratory and the Department of Electrical and Computer Engineering at McMaster University in 1983 as a graduate student. He was awarded an Ontario Graduate Scholarship for the academic years 1985/86 and 1986/87. He was a teaching assistant in the Department of Electrical and Computer Engineering, McMaster University, from 1984 to 1987 and a postdoctorate fellow there from 1987 to 1988. Presently he is a research engineer with Optimization Systems Associates Inc., Dundas, Ontario, Canada. In 1989, he also became an Assistant Professor (part-time) in the Department of Electrical and Computer Engineering at McMaster University. His professional interests include all aspects of circuit CAD with emphasis on large-scale simulation and optimization, design and modeling of nonlinear microwave circuits, statistical modeling, sensitivity analysis, and the diagnosis and tuning of analog circuits.

Dr. Zhang was a contributor to *Analog Methods for Computer-Aided Analysis and Diagnosis* (Marcel Dekker, 1988).

Jian Song was born in Anshan, Liaoning, China, on May 5, 1956. He received the B.Eng. degree and the M.Eng. degree from Chongqing University, Chongqing, China, in 1982 and 1984, respectively, both in electrical engineering.

In 1984 he joined the Electrical Engineering Department, Chongqing University, Chongqing, China, where he was involved in research in the area of circuit theory and applications. Since 1986, he has been with the Simulation Optimization Systems Research Laboratory and the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada, as a graduate student working toward the Ph.D. degree. His research interests include general circuit theory, statistical circuit design, yield optimization techniques, device statistical modeling, optimization methods, and large-scale numerical techniques.



Radoslaw M. Biernacki (M'86–SM'86) was born in Warsaw, Poland. He received the Ph.D. degree from the Technical University of Warsaw, Warsaw, Poland, in 1976.

He became a Research and Teaching Assistant in 1969 and an Assistant Professor in 1976 at the Institute of Electronics Fundamentals, Technical University of Warsaw, Warsaw, Poland. From 1978 to 1980 he was on leave with the Research Group on Simulation, Optimization and Control and with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, as a Post-Doctorate Fellow. From 1984 to 1986 he was a Visiting Associate Professor at Texas A&M University, College Station, TX. He joined Optimization Systems Associates, Inc., Dundas, Ontario, Canada, in 1986 as Senior Research Engineer. In 1988 he also became a Professor (part-time) in the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada. His research interests include system theory, optimization and numerical methods, computer-aided design of integrated circuits, and control systems.

Dr. Biernacki has more than 50 publications and has received a number of prizes for his research and teaching activities.