FAST GRADIENT BASED YIELD OPTIMIZATION OF NONLINEAR CIRCUITS

J.W. Bandler, Q.J. Zhang, J. Song and R.M. Biernacki

SOS-90-7-R

February 1990

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 *I*ntegrated *G*radient *A*pproximation *T*echnique (*IGAT*) and the *F*easible *A*djoint *S*ensitivity *T*echnique (*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 *P*erturbation *A*pproximate *S*ensitivity *T*echnique (*PAST*) on the one extreme and the theoretical *E*xact *A*djoint *S*ensitivity *T*echnique (*EAST*) on the other. *FAST*, linking state-of-the-art optimization and efficient HB 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 6 optimizable variables including input power and bias conditions, and 34 statistical parameters. Using either *IGAT* or *FAST* yield is driven from 25% to 67%. *FAST* exhibits superior efficiency.

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

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 fifteen 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.

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 *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 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 *E*xact *A*djoint *S*ensitivity *T*echnique (*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 an *I*ntegrated *G*radient *A*pproximation *T*echnique (*IGAT*) presented by Bandler et al. [17], which is adapted here to the needs of yield optimization. The other is the *F*easible *A*djoint *S*ensitivity *T*echnique (*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 25% to 67%.

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

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

if it is an upper specification, or

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

in the case of lower specifications, where

$$\mathbf{h} = [\ 0\ 1\ 2\ .\ .\ .\ H\ ]^T \tag{2}$$

is the harmonic index vector, 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, ϕ^0, along with a statistical distribution of parameters. Parameters in ϕ^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 ϕ^i, can be written as

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

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

The response of each outcome, denoted by

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

is calculated after solving the HB equations [19]

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

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

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

or as

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

We assemble all errors for the outcome ϕ^i into one vector $\mathbf{e}^i$. If all entries of this vector are nonpositive, the outcome ϕ^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 ϕ^0. 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

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

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 e^i [3],

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

where

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

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

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

where

$$I = \{i \;|\, v^i > 0\} \tag{12}$$

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

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

then function $U(\phi^0)$ 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 general, the functions $e_j(\phi^i)$ are differentiable, but the functions v^i of (9) may not be. The problem implied by (16) is better posed than that of (11). 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 made 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 the Fig. 1.

III. INTEGRATED GRADIENT APPROXIMATION TECHNIQUES (*IGAT*)

We review the perturbation approximate sensitivity technique (*PAST*). Then *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 $f(\boldsymbol{\phi})$ to denote a generic function.

A. Approximating Derivatives by PAST

The first-order derivative of $f(\boldsymbol{\phi})$ with respective to the *k*th variable can be estimated by

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

where $\boldsymbol{\phi} + \Delta\phi_k \mathbf{u}_k$ denotes the perturbation of the *k*th variable and where $\Delta\phi_k$ is the perturbation length and $\mathbf{u}_k$ is a column vector which has 1 in the *k*th position and zeros elsewhere. An approximation to the gradient, $\nabla f(\boldsymbol{\phi})$, 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,

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

where $\boldsymbol{\phi}_{old}$ and $\boldsymbol{\phi}_{new}$ are two different points and $\Delta\boldsymbol{\phi} = \boldsymbol{\phi}_{new} - \boldsymbol{\phi}_{old}$. If $\boldsymbol{\phi}_{old}$ and $\boldsymbol{\phi}_{new}$ are iterates of optimization, $f(\boldsymbol{\phi}_{old})$ and $f(\boldsymbol{\phi}_{new})$ 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* (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 3 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, ϕ^0, are design variables, all perturbations are made to ϕ^0 in the initialization and reinitialization steps using *PAST* (17). When ϕ^0 is perturbed to $\phi^0 + \Delta\phi^0_k \mathbf{u}_k$, denoted by $\phi^0_{k,pert}$ for short, outcomes should be regenerated from $\phi^0_{k,pert}$ in order to get perturbed circuit responses. These outcomes are denoted by $\phi^i_{k,pert}$. Then from (17), the approximate derivative of the response $f(\phi^i)$ is defined as

$$\frac{\partial f(\phi^i)}{\partial \phi^0_k} \approx \frac{f(\phi^i_{k,pert}) - f(\phi^i)}{\Delta\phi^0_k}. \tag{19}$$

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

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

IV. FEASIBLE ADJOINT SENSITIVITY TECHNIQUE (*FAST*)

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

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

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

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

Bandler, Zhang and Biernacki [13,14] have proposed the *E*xact *A*djoint *S*ensitivity *T*echnique (*EAST*). The sensitivity expressions for various elements have been derived and listed [14].

Here, we propose the *F*easible *A*djoint *S*ensitivity *T*echnique (*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

$$\overline{V}_{out} = [\mathbf{a}^T \;\; \mathbf{b}^T] \begin{bmatrix} \overline{\mathbf{V}} \\ \overline{\mathbf{V}}_s \end{bmatrix} = \mathbf{c}^T \begin{bmatrix} \overline{\mathbf{V}} \\ \overline{\mathbf{V}}_s \end{bmatrix}, \tag{23}$$

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

$$\mathbf{c} = [\mathbf{a}^T \;\; \mathbf{b}^T]^T \tag{24}$$

is a linear transfer vector linking the output voltage with $\overline{\mathbf{V}}_s$ and $\overline{\mathbf{V}}$. $\overline{\mathbf{V}}$ and $\mathbf{c}$ are functions of $\boldsymbol{\phi}$. $\overline{\mathbf{V}}_s$ can also be a function of $\boldsymbol{\phi}$ if we want to change $\overline{\mathbf{V}}_s$ 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 $\overline{V}_{out}$ w.r.t. ϕ_k can be calculated as

$$\frac{\Delta \overline{V}_{out}}{\Delta \phi_k} \approx \frac{\Delta \mathbf{c}^T}{\Delta \phi_k} \begin{bmatrix} \overline{\mathbf{V}} \\ \overline{\mathbf{V}}_s \end{bmatrix} + \mathbf{a}^T \frac{\Delta \overline{\mathbf{V}}}{\Delta \phi_k} + \mathbf{b}^T \frac{\Delta \overline{\mathbf{V}}_s}{\Delta \phi_k}, \tag{25}$$

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

$$\mathbf{J}^T \hat{\overline{\mathbf{V}}} = \mathbf{a}, \tag{26}$$

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

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

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

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

for a small $\Delta\phi_k$.

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

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

This formula is much easier to implement than the corresponding formula for *EAST* [13, 14]. The function $\overline{\mathbf{F}}(\boldsymbol{\phi} + \Delta\phi_k \mathbf{u}_k, \overline{\mathbf{V}})$ 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 $\overline{\mathbf{V}}$ and $\overline{\mathbf{V}}_s$ are also available from the HB simulation. The perturbed vectors $\mathbf{a}(\boldsymbol{\phi} + \Delta\phi_k \mathbf{u}_k)$ and $\mathbf{b}(\boldsymbol{\phi} + \Delta\phi_k \mathbf{u}_k)$ can be easily calculated since they involve the linear subnetwork only. Finally, the perturbed excitations $\overline{\mathbf{V}}_s(\boldsymbol{\phi} + \Delta\phi_k \mathbf{u}_k)$ 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 $\overline{\mathbf{V}}_{out}$ w.r.t. ϕ_k can be computed as

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

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. 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 *i*th outcome due to the perturbation is calculated by

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

where

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

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

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 10 design variables, instead of completely solving 10 nonlinear circuits, we only evaluate 10 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 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.

VI. YIELD OPTIMIZATION OF A FREQUENCY DOUBLER

Consider the FET frequency doubler shown in Fig. 2 [20]. It consists of a common-source FET with a lumped input matching network and a microstrip output 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_{GB} and V_{DB} 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, 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 19 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 to be 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, then these error functions with their multipliers defined in (13) are fed into the one-sided ℓ_1 optimizer. *IGAT* and *FAST* are implemented to provide gradients. *IGAT* calculates approximate sensitivities of the output powers, which are later converted to the gradients of the conversion gain and spectral purity.

The starting point for yield optimization is the solution of the conventional nominal design w.r.t. the same specifications, using L_1, l_1 and l_2 as optimization variables. The initial yield based on 500 outcomes is 24.8%.

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. For the first and second phases, two different sets of 50 statistically generated outcomes are used.

Using *IGAT*, the first and second phases reach 55.8% and 67.6% yields, respectively. The two phases use 20 optimization iterations and 62 function evaluations, equivalent to 3100 circuit simulations. For *FAST*, the first phase uses only 8 function evaluations and gradient calculations to give 67.2% yield. The second phase slightly increases the estimated yield to 67.4%, 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 first phase of the *FAST* approach is much less than the total CPU time used by the *IGAT* approach. Although *IGAT* is slower than *FAST*, it is very robust in terms of the final yield reached.

Figs. 3, 4 and 5 show histograms of the conversion gain before and after yield optimization. Fig. 3 is the conversion gain distribution of 500 outcomes before yield optimization. The histograms in Figs. 4 and 5 are based on solutions using *IGAT* and *FAST*, respectively. The improvement of circuit performance is very clearly illustrated by the histograms. Before yield optimization, the center of the distribution is on the left-hand side of the design specification of 2.5 dB, indicating that most outcomes are unacceptable. After yield optimization, the center of the distribution is shifted to the right-hand side of the 2.5 dB specification. Most outcomes then satisfy the specification.

Design with *PAST* has also been carried out. To reach the same yield level, it uses a total of 20 optimization iterations in two phases, equivalent to 140 function evaluations or 7000 circuit simulations. The total CPU time is 31 minutes, which is 2.5 and 4 times the CPU times required by *IGAT* and *FAST*, respectively.

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 advantage 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," *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, 1988, pp. 424-443.

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

[5] R. S. Soin and R. Spence, "Statistical exploration approach to design centering," *Proc. Inst. Elec. Eng.*, vol. 127, pt. G., 1980, pp. 260-269.

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

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

[8] M. A. Styblinski and A. Ruszczynski, "Stochastic approximation approach to statistical circuit design," *Electron. Lett.*, vol. 19, no. 8, 1980, pp. 300-302.

[9] E. Polak and A. L. Sangiovanni-Vincentelli, "Theoretical and computational aspects of the optimal design centering, tolerancing, and tuning problem," *IEEE Trans. Circuits and Systems*, vol. CAS-26, 1979, pp. 795-813.

[10] K. Singhal and J.F. Pinel, "Statistical design centering and tolerancing using parametric sampling," *IEEE Trans. Circuits and Systems*, vol. CAS-28, 1981, pp. 692-701.

[11] R.M. Biernacki, J.W. Bandler, J. Song and Q.J. Zhang, "Efficient quadratic approximation for statistical design," *IEEE Trans. Circuits and Systems*, vol. CAS-36, 1989, pp. 1449-1454.

[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, 1983, pp. 762-770.

[13] J.W. Bandler, Q.J. Zhang and R.M. Biernacki, "A unified framework for harmonic balance simulation and sensitivity analysis," *IEEE 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, 1988, pp. 1661-1669.

[15] J.W. Bandler, Q.J. Zhang, J. Song and R.M. Biernacki, "Yield optimization of nonlinear circuits with statistically characterized devices," *IEEE 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," *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, 1988, pp. 444-454.

[18] J.W. Bandler, Q.J. Zhang and R.M. Biernacki, "Practical, high-speed gradient computation for harmonic balance simulators," *IEEE 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, 1987, pp. 521-535.

[20] *Microwave Harmonica User's Manual*, Compact Software Inc., Paterson, NJ 07504, 1987.

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

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

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(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 [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.

TABLE III

YIELD OPTIMIZATION OF THE FREQUENCY DOUBLER USING *IGAT*

Variable	Starting Point	Nominal Design	Solution I	Solution II
P_{IN}(W)	$2.0000{\times}10^{-3*}$	$2.0000{\times}10^{-3}$	$2.1442{\times}10^{-3}$	$1.7507{\times}10^{-3}$
V_{GB}(V)	-1.9060*	-1.9060	-1.9062	-1.7890
V_{DB}(V)	5.0000*	5.0000	5.0004	5.4504
L_1(nH)	1.0000	5.4620	5.4623	5.4665
l_1(m)	$1.0000{\times}10^{-3}$	$1.4828{\times}10^{-3}$	$1.7027{\times}10^{-3}$	$1.7039{\times}10^{-3}$
l_2(m)	$5.0000{\times}10^{-3}$	$5.7705{\times}10^{-3}$	$5.7573{\times}10^{-3}$	$5.7629{\times}10^{-3}$
Yield		24.8%	55.8%	67.6%
No. of Optimization Iterations			9	11
No. of Function Evaluations			27	35
CPU Time (Multiflow Trace 14/300)			5.2min	7.2min

* Not considered as variables in the nominal design.

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)	$2.0000\times10^{-3*}$	2.0000×10^{-3}	1.7581×10^{-3}	1.9107×10^{-3}
V_{GB}(V)	-1.9060^{*}	-1.9060	-1.8578	-1.8576
V_{DB}(V)	5.0000^{*}	5.0000	5.5000	5.5000
L_1(nH)	1.0000	5.4620	5.4635	5.4637
l_1(m)	1.0000×10^{-3}	1.4828×10^{-3}	1.6707×10^{-3}	1.7623×10^{-3}
l_2(m)	5.0000×10^{-3}	5.7705×10^{-3}	5.7642×10^{-3}	5.7751×10^{-3}
Yield		24.8%	67.2%	67.4%
No. of Optimization Iterations			8	10
No. of Function Evaluations and Sensitivity Analyses			8	10
CPU Time (Multiflow Trace 14/300)			3.4min	4.4min

* Not considered as variables in the nominal design.

The yield is estimated from 500 outcomes.

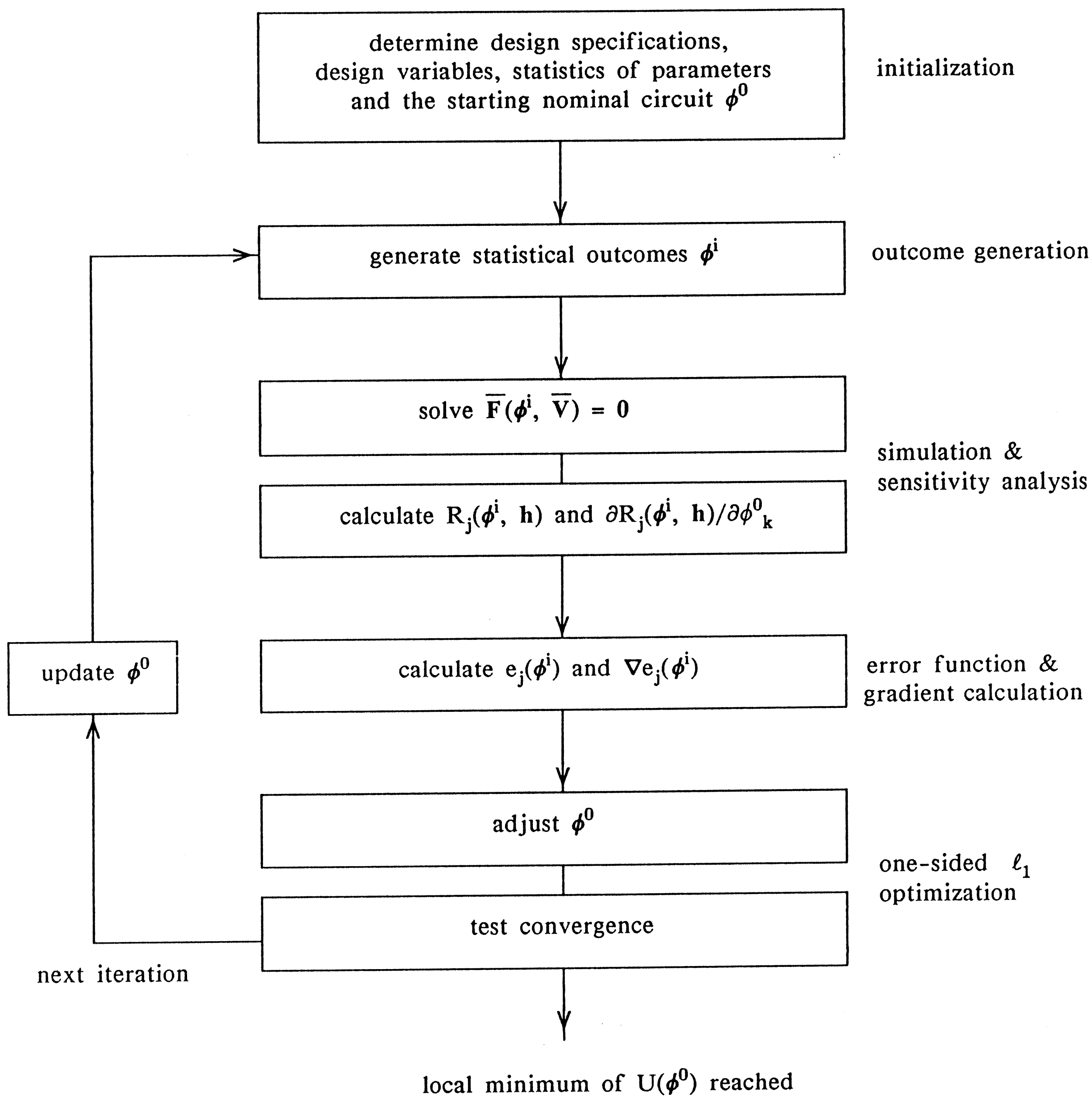

Fig. 1 Flowchart for yield optimization.

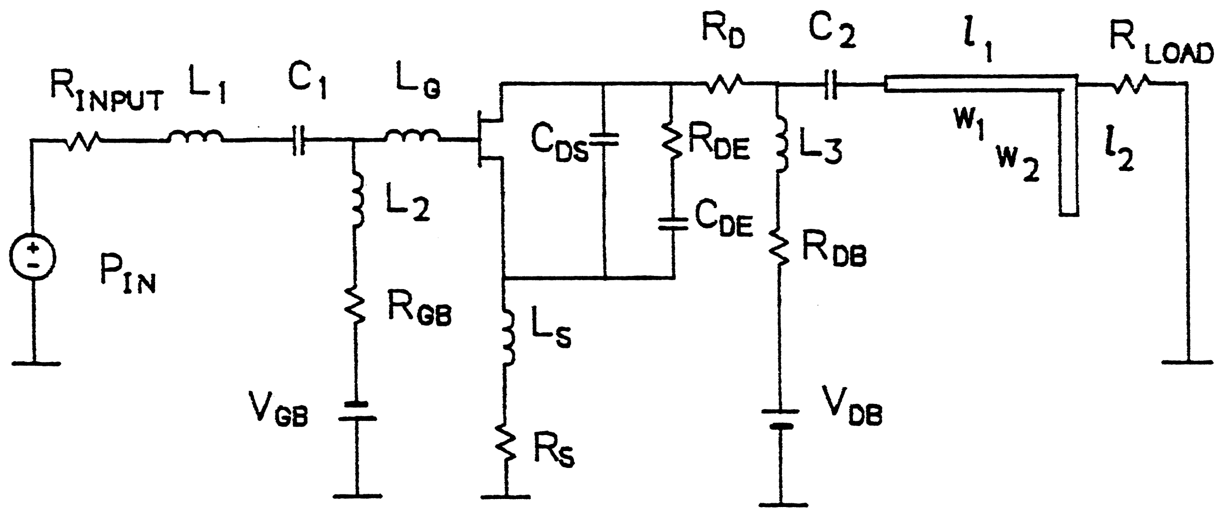

Fig. 2 Circuit diagram of the FET microwave frequency doubler. The nominal values for non-optimizable variables are: L_2 = 15nH, L_3 = 15nH, C_1 = 20pF, C_2 = 20pF, $w_1 = 0.1\times10^{-3}$m, $w_2 = 0.635\times10^{-3}$m, $R_{LOAD} = R_{INPUT}$ = 50Ω, and $R_{GB} = R_{DB}$ = 10Ω.

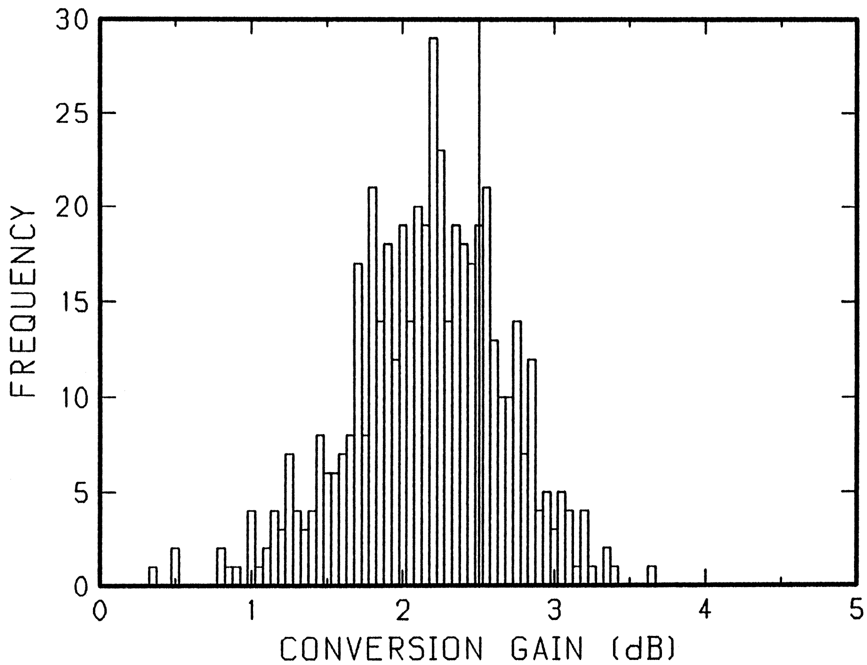

Fig. 3 Histogram of conversion gains of the frequency doubler based on 500 statistical outcomes at the starting point for yield optimization. The specification is shown by a vertical line.

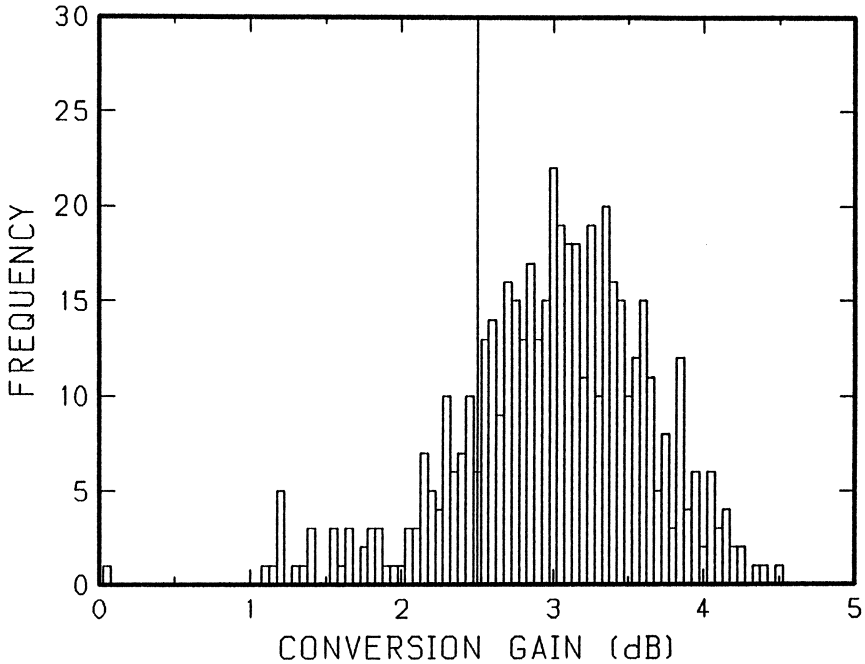

Fig. 4 Histogram of conversion gains of the frequency doubler based on 500 statistical outcomes at the second phase solution using *IGAT*. The specification is shown by a vertical line.