Yield Optimization Using a GaAs Process Simulator Coupled to a Physical Device Model

Dan E. Stoneking, *Member, IEEE,* Griff L. Bilbro, Paul A. Gilmore, Robert J. Trew, *Fellow, IEEE,* and C. T. Kelley

(Invited Paper)

***Abstract*—A physics based large-signal GaAs MESFET model and circuit simulator has been developed to predict and optimize the yield of GaAs MESFET designs before fabrication. Device acceptance criteria include both small- and large-signal RF operating characteristics such as small-signal gain, maximum power added efficiency, and output power at 1 dB gain compression. Channel doping details are described directly from processing specifications for parameters such as material deposition, ion implantation, and implant annealing. Monte Carlo techniques are used to estimate yield when disturbances in the physical parameters are modeled as multivariate Gaussian distributions. The yield estimator is integrated with an optimizer so that a design can be *centered* for maximum yield in the presence of process disturbances.**

I. Introduction

COMPUTER-AIDED DESIGN (CAD) has helped to improve process yields and the average performance of monolithic microwave integrated circuits (MMIC's) through the application of statistical circuit design techniques. A principal statistical circuit design technique is yield optimization.

The yield optimization problem can be formulated two ways:

$$\max_x \left\{ Y(x) = \int_{R_A} p(v)\, dv \right\}$$

or

$$\max_x \left\{ Y(x) = \int_{-\infty}^{\infty} p(v)\, \phi(x + v)\, dv \right\}$$

where $x \in R^n$, $p(v)$ is the parameter disturbance probability density function, and R_A is the acceptability region. The acceptance function $\phi(x + v) = 1$ if $(x + v) \in R_A$. Otherwise, $\phi(x + v) = 0$. The first formulation requires approximating the region of acceptability and leads to region of acceptability approaches. The second formulation is usually solved with Monte Carlo techniques.

A number of different authors proposed region of acceptability approaches to the yield optimization problem. Scott and Walker [1] and Leung and Spence [2] pursued the regionalization method. Regionalization does a direct search on the space to determine the acceptability region. The simplicial approach was used by Director *et al.* [3]–[5]. The simplicial approach approximates the region of acceptability with a polyhedron formed by points on the boundary. The design center is then the center of a hypersphere which is contained within the polyhedron. Abdel-Malek and Bandler [6]–[8] approximated the acceptability region with hyperboxes and linear cuts. The yield is estimated as the hypervolume of the region of acceptability divided by the hypervolume of the tolerance box. Ellipsoids of decreasing volume were proposed by Abdel-Malek and Hassan [9]. This method approximates the region of acceptability with an ellipsoid which is determined by decreasing the volume, moving the center, and changing the shape of a starting ellipsoid. The design center is then center of the final ellipsoid.

Various authors sought to solve the yield optimization problem with Monte Carlo techniques. Soin and Spence [10] advocated a center of gravity method. N Monte Carlo samples are taken. The centers of gravity of the passed and failed points are determined. The design center is then located by a line search along the line joining the centers of gravity. Stochastic optimization was used by Styblinski and Ruszczynski [11] and Kjellström and Taxèn [12]. In this approach, small sample Monte Carlo yield estimates are stochastically optimized. Signhal and Pinel [13] introduced parametric sampling. The parametric sampling technique reuses previous samples when forming a Monte Carlo yield estimate. The control variate technique [14] for reducing the Monte Carlo noise in the yield estimate was used by Hocevar *et al.* [15] and Soin and Rankin [16], [17]. Biernacki *et al.* [18] used efficient quadratic approximation. This technique approximates the circuit response with a multidimensional quadratic function which reduces the cost of the Monte Carlo samples. The approximation is generated from less than the minimum number of basis points by a maximally flat interpolation technique. Ban-

Manuscript received August 20, 1991; revised February 13, 1992. This work was supported by an ARO Fellowship award to D. E. Stoneking, a Phase 3 MIMIC Program award No. DAAL01-89-K-0906, NSF Grants No. DMS-8900410, and No. DMS-9024622, and AFOSR Grants No. AFOSR-ISSA-890044 and No. AFOSR-FQ8671-9101094.

D. E. Stoneking is with M/A-COM Corporation, Lowell, MA.

G. L. Bilbro and R. J. Trew are with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27695-7911.

P. A. Gilmore and C. T. Kelley are with the Department of Mathematics, North Carolina State University, Raleigh, NC 27695-7911.

IEEE Log Number 9108313.

dler *et al.* applied minimization of the generalized l_p norm to yield optimization [19]. Various approaches to gradient calculation in the context of generalized l_p norms are discussed by Bandler *et al.* [20]. The l_p formulation is used by Bandler *et al.* [21] in conjunction with the Khatibzadeh–Trew MESFET model [22], [23] to optimize an *X*-band amplifier.

Reviews of yield optimization and statistical circuit design techniques are available. Some of this review material appears in books by various editors and authors [5], [24]–[27]. Review papers on the topic include [19], [28].

Most previous yield optimization work is based upon the use of equivalent circuit models for active devices. This dependence ultimately limits present generation CAD since equivalent circuit models do not *predict* operation so much as compactly represent measurements on fabricated devices. Equivalent circuit elements do not naturally capture important nonlinearities and interrelationships of the physical entities they represent. Also, determining improved device design parameters from equivalent circuit elements is difficult because direct correspondence between physical parameters, equivalent circuit elements, and RF performance is not easily established. Equivalent circuit models do not scale well to operating regions exterior to those in which the model was directly calibrated.

In this work, a large-signal GaAs MESFET simulator for yield estimation and optimization is presented that does not rely on equivalent circuit techniques. The MESFET model in the simulator is based upon device physics with the advantage that a MIMIC can be simulated from process data all the way through to RF circuit performance. The integrated simulator allows a device design to be optimized based upon a desired RF performance specification. Small- and large-signal power amplifier performance measures are used for the yield pass-fall criterion. The yield optimizer's variables are physical parameters such as gate dimensions, channel donor distribution specifications, dc bias voltages, and material parameters. A Monte Carlo method is used to predict the process yield of a nominal MESFET design based upon a single performance measures' variation. The Monte Carlo yield estimate is then optimized using a quasi-Newton method. The quasi-Newton method is deterministic and is tolerant of the inherent noise in the yield estimate.

The resulting simulator has proven to be accurate under both class A and B operating conditions. Since simulations can be performed before fabrication, significant time, effort, and expense can be saved in the development of advanced MMIC's.

Section II describes the components of the simulator pertinent to yield optimization. The results of yield optimization experiments on an ion implanted device, a buried channel device, and a uniform channel device are presented in Section III.

II. Simulator Components

Computer optimization of the yield of a device using physical variables requires an algorithm that integrates a physical model for the device, a method for computing performance measures, a yield estimation method, and an optimization algorithm. Pertinent aspects of these components are discussed in this section.

A. Device Model

The device model used in this work [22], [23] is based upon an efficient, analytic solution of the basic semiconductor device equations. The device model is quasi-static and is based upon an analytic formulation of the charge dipole domain within the conducting channel. This formulation has, in turn, resulted in the derivation of a self-consistent large-signal analytic MESFET model that can easily be implemented in microwave CAD simulators. The time domain analytic device model is interfaced with an RF circuit by means of the harmonic balance method to produce an efficient and accurate nonlinear device/circuit microwave simulator.

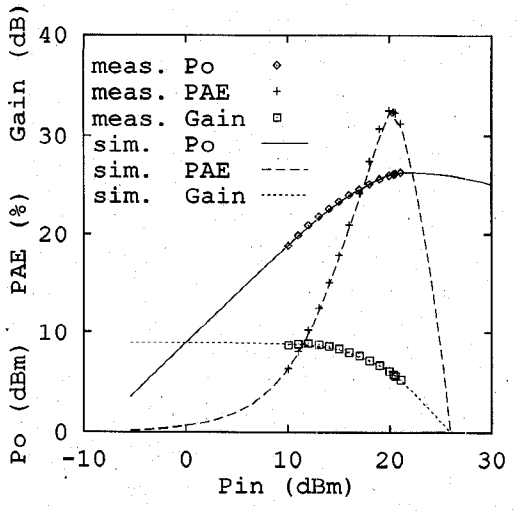
A recent enhancement of the MESFET model for use with ion-implanted devices is the incorporation of SUPREM 3.5 [29]. SUPREM 3.5 calculates the MESFET channel donor distribution resulting from process sequences which may include material deposition and etching, dopant diffusion, ion implantation, or implant annealing. With this enhancement, a design can be specified in terms of process variables instead of donor distributions.

The model, when embedded in a harmonic balance simulator, has been verified to accurately predict the dc and RF behavior of several commercial devices [22]. For example, Fig. 1 demonstrates the measured and simulated RF performance for a commercial 0.5 μm gate length ion-implanted power GaAs MESFET at 10 GHz. The device was operated class A at a drain bias of $V_{dd} = -6.5$ V. The physical parameters used in the simulation are indicated. The generator and load impedances for the fundamental, second, and third harmonics are also indicated. The excellent agreement between the measured and simulated data results from inclusion of the major saturation mechanisms in a GaAs MESFET. These mechanisms are forward conduction and reverse breakdown of the gate electrode [30], [31]. Models for these mechanisms suitable for integration into the simulator have been developed [32].

The MESFET model is embedded in a linear circuit as shown in Fig. 2. The harmonic balance algorithm solves for the voltages and currents in the circuit when Vgen has some specified value. Using this system, performance measures applicable to MESFET power amplifiers are computed.

B. Performance Measures

The performance measures fall into three different classes: small-signal transducer gain, measures associated with gain compression levels, and measures associated with maximum power added efficiency.



(a)

MESFET Parameter Values	
W_g	1126 μm
L_g	0.509 μm
N_{max}	2.81×10^{17}cm^{-3}
Range	0.104 μm
Staggle	0.050 μm
t_{chnl}	0.471 μm
v_{sat}	1.299×10^{7}cm/s
μ	4000 cm/V · s
E_{crit}	3.246 kV/cm
R_{gg}	0.91 Ω
R_{dd}	1.25 Ω
R_{ss}	1.68 Ω
L_{gg}	0.11 nH
L_{dd}	0.17 nH
L_{ss}	0.001 nH
V_{dsbd}	11.20 V
BBS	-0.5 V/V
R_{gdbd}	24.73 Ω
R_{gsfb}	1.0 Ω

Circuit Impedances	
$Z_{gen,1}$	56 + j34Ω
$Z_{gen,2}$	50 + j50Ω
$Z_{gen,3}$	40 + j50Ω
$Z_{load,1}$	55 + j13Ω
$Z_{load,2}$	12 + j9Ω
$Z_{load,3}$	100 + j40Ω

Bias	
V_{GG}	-1.87 V
V_{DD}	-6.50 V

(b)

Fig. 1. Comparison of measured and TEFLON simulated data. (a) Three RF performance characteristics of a MESFET (operational gain, output power delivered to the load, and operational power added efficiency). (b) MESFET, circuit, and bias parameter values during simulation.

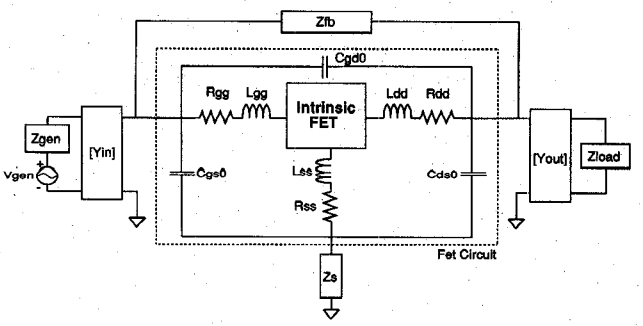


Fig. 2. The complete TEFLON RF circuit schematic. The intrinsic FET block represents the physics based Khatibzadeh–Trew model.

The small-signal transducer gain is given by

$$G_{\text{ave}} = \frac{1}{4}\sum_{k=1}^{4}\left(P^k_{l,\text{dB}} - P^k_{g,\text{dB}}\right)$$

where $P_{g,\text{dB}}$ is the RF power in dBm available from the generator Vgen and $P_{l,\text{dB}}$ is the RF power in dBm delivered to the load. The four values of $P_{g,\text{dB}}$ are specified by the user, and they must be in the small-signal regime. The small-signal transducer gain has no other associated performance measures.

The gain compression measures result from solution of

$$\left(P_{l,\text{dB}}(P_{g,\text{dB}}) - P_{g,\text{dB}}\right) - \left(G_{\text{ave}} - G_C\right) = 0 \qquad P_{g,\text{dB}} \in R$$

where G_C is a gain compression level in dB and is user selectable as 1, 3, or 6 dB. The Van Wijingaarden–Dekker–Brent root finding algorithm [33] is used to solve the above equation. Once the root of the equation is found, the available performance measures are $P_{g,\text{dB}}$, and $P_{l,\text{dB}}$ at the solution.

The final class of performance measures result from maximizing the expression for the power added efficiency,

$$\text{PAE} = \frac{P_l(P_g) - P_g}{P_{\text{dc}}(P_g)} \qquad P_g \in R$$

where P_l is the RF power in mW delivered to the load, P_g is the RF power in mW available from the generator, and P_{dc} is the dc power in mW delivered to the transistor. Brent's method [33] is used for this 1-D optimization. The resulting performance measures are the PAE and P_g, P_l, and P_l/P_g at the maximum.

Optimizations using these performance measures reveal the need for additional constraints beyond an optimization variable hyperbox. These constraints are implemented as penalty functions on the measures. A penalty function is invoked under the following conditions:

1. the device pinch-off voltage V_{po} is not within bounds,
2. the dc bias is not within bounds,
3. the device power gain is less than a minimum value, and
4. the power flow at the 1st harmonic is into the RF source (i.e., the device is oscillating).

C. Yield Estimation Method

Small but uncontrollable disturbances in the fabrication process result in devices with geometries and doping profiles that deviate somewhat from nominal values. These variations in primary process parameters are statistically simpler than the derivative variations of parameters for a corresponding equivalent circuit. For example, gate width and length are practically independent and both are uncorrelated with ion implant dose or energy. A problem with equivalent circuit based yield approaches is that variations in equivalent circuit parameter values such as g_m, R_i, C_{gs}, C_{dg}, C_{ds}, and g_{ds} correlate significantly with each

other [34]–[36]. Moreover, second order correlations (even large ones) may not suffice to characterize variations in equivalent circuit parameters [37].

In the approach presented here disturbances in physical parameters are easily characterized by a second order statistical model, a multivariate Gaussian. A multivariate Gaussian is specified by a mean vector and a covariance matrix. The mean vector is simply the nominal device design. The covariance matrix models the variances and covariances between physical parameters.

Given the covariance matrix C for the parameter disturbance probability density function $p(v)$, the yield at some nominal design, x, is estimated by the following Monte Carlo algorithm:

1. Input mean or nominal design x,
2. Generate a set of N disturbances v_i from the distribution $p(v)$ [33], [38],
3. Approximate the yield as

$$Y(x) \approx \frac{1}{N} \sum_i \phi(x + v_i).$$

In a Monte Carlo yield calculation, many devices are evaluated in the vicinity of a nominal design. In regions where the performance measure is relatively smooth, it is possible to use nearby points without incurring unacceptable estimation errors. The batch size of a Monte Carlo calculation sets a limit on expected precision which can be related to derivatives of the performance measure to determine when it is appropriate to replace a numerically intensive performance measure calculation with a previously simulated result.

To reduce run times a simple memory system which maintains a binary tree database of previous simulations is incorporated. When a new design is to be evaluated, the tree is searched for a similar design. If a previously calculated design and the desired design are within a user specified neighborhood of each other, the stored performance measure is returned. Otherwise, the device model is called and the simulation result is recorded in the tree. This memory system gives a piecewise constant interpolation to the performance measure.

The neighborhood is a single hyperbox from an equal volume grid of the optimization constraint hyperbox. In each dimension, the edge of a neighborhood hyperbox is 2^{-M} times the corresponding edge of the optimization constraining hyperbox. M is a user specified integer that determines the factional neighborhood hyperbox volume. Thus, the neighborhood hyperbox volume is 2^{-MD} times the constraining hyperbox volume. D is the number of dimensions in the space. For the yield optimization results presented in Section III, M was set to 4.

This memory system will reduce the number of calls by varying amounts dependent upon the particular device design, the degree of nonlinearity, the number of disturbance parameters, and the extent of previous data. For the yield optimization results presented in Section III, the number of calls is reduced by 17%. As the memory grows larger the number of calls is further reduced. Reductions as high as 65% are observed for large memories.

Earlier work using linear interpolations of the memory data provided greater reductions of device model calls. Reductions as high as 90% were observed. However, linear interpolations provide accurate approximations only if the data being interpolated is relatively smooth. When linearly interpolating in the vicinity of a step penalty function, large approximation errors result. For this reason, the piecewise constant interpolation approach was used.

D. Mathematical Optimizer

The final component of the yield optimization algorithm is the mathematical optimizer. A primary consideration in developing the optimizer was the inherent noise of the Monte Carlo yield estimate. The noise is approximately given by [14], [28]

$$\text{var}(Y) \simeq \frac{Y(1 - Y)}{N},$$

where Y is the estimate of the yield and N is the simulated batch size. This noise term led to the development of the optimization algorithm discussed below. Using this optimizer, batch sizes of 100 are necessary for successful performance although batch sizes as small as 50 can be useful for preliminary experiments.

The optimizer is a projected quasi-Newton algorithm that utilizes a decreasing sequence of finite difference steps (scales) to approximate the yield gradient. This decreasing sequence of scales permits this deterministic algorithm to deal with the Monte Carlo noise in the yield estimate. The algorithm uses an efficiently calculated approximation to the Hessian and a line search algorithm that gives the code global convergence properties.

A mathematical description of the problem of constrained optimization goes as follows:

$$\min_Q f: \; Q \to R \text{ where}$$

$$Q = \{x \in R^n | l_i \le x_i \le u_i, \; i = 1, \cdots, n\}.$$

The set of points Q is called the hyperbox, and f equals $-Y$.

The optimization code is a variation of Newton's method for unconstrained optimization described later in this section.

During yield optimization the function f to be minimized is of the form: $f(x) = \hat{f}(x) + \Delta f(x)$ where $\hat{f}(x)$ is the negative of the true yield, $\bar{\phi}$, and $\Delta f(x)$ is Monte Carlo noise. This noise creates a large number of local minima. $\bar{\phi}$ may itself have local minima that are not the global minima. This effect may be a consequence of physical phenomena or of error in the harmonic balance simulation used to compute $\hat{f}$. These local minima will trap most gradient based algorithms. However, by utilizing a sequence of scales to approximate the gradient these local minima can be avoided, and the global minima can be found up to the level of noise.

The novel feature of the optimization algorithm is this sequence of scales. The first elements in the sequence of finite difference steps are fairly large, approximately half the length of the hyperbox. The approximate gradient obtained using these scales gives global information about the problem and allows the sequence of quasi-Newton steps to avoid local minima. As the sequence continues, the length of the scales decreases so that the approximate gradient gives more local information. As the algorithm steps through the sequence of scales the code becomes more of a local algorithm having the fast local convergence properties of quasi-Newton methods. Thus, the sequence of scales allows the algorithm to avoid the local minima generated by the noise and also have fast local convergence properties.

When computing the yield gradient with different formulas, an important consideration is minimizing the variance or noise in the gradient estimate. The method of common random numbers [14] was applied to reduce the gradient variance. This method requires that during gradient computations every evaluation of f be computed with the same sequence of random disturbances.

In Newton's method the Hessian must be computed at each point to calculate the step. Because analytic second derivatives are not available and because the cost of function evaluations is too high to use a difference Hessian, an approximation to the Hessian is used. The most successful approximation yet applied is the SR1 (symmetric rank one). This approximation is given by the following formula:

$$S^c = S^- + \frac{r^- r^{-T}}{r^{-T} s^-}.$$

Where S^c and S^- are the current and previous SR1 approximations to the Hessian respectively, $s^- = x^c - x^-$, where x^c and x^- are the current and the previous points in the sequence defined by the optimization algorithm respectively, and, $r^- = y^- - S^- s^-$, where $y^- = \nabla f(x^c) - \nabla f(x^-)$. Note that calculating the SR1 approximation to the Hessian does not involve taking any extra function evaluations and hence is far cheaper than using a difference Hessian.

Some notation that will be used later is now defined. The projection $v^\wedge$ of a vector v onto the hyperbox Q is given by the following formula

$$v_i^\wedge = \begin{cases} u_i, & \text{if } v_i > u_i \\ v_i, & \text{if } l_i \le v_i \le u_i, \\ l_i, & \text{if } v_i < l_i \end{cases}$$

A description of the algorithm is as follows:

1. Given x^c, h, S^c, calculate $f(x^c)$, $\nabla f(x^c)$, and $y = (x^c - \nabla f(x^c))^\wedge$. If
$$\|x^c - y\| \le \frac{\text{minscal}}{2h}$$
set $h = 0.5h$, and $S^c = I$. If $h <$ *minscal* or $f > f_{\max}$ terminate, otherwise return to 1. Where *minscal* is the minimal scale used and $f_{\max}$ is the maximum possible value for f. *Minscal* is an estimate of the minimum distance between points in the hyperbox for which an appreciable difference in the value of f can be detected. For yield optimization $f \le 1$, however, $f_{\max}$ is set to 0.98 for this application.
2. Otherwise update S^c as described above. Check if S^c is positive definite. If it is not, set $S^c = I$. In our code positive definiteness of the matrix is checked in the Linpack code DCHDC [39], a double precision Cholesky decomposition routine. Setting $S^c = I$ if the approximate Hessian is not positive definite guarantees that the step is always a descent direction.
3. Solve $S^c p = \nabla f(x^c)$. If $\|p\| \le h$ then set
$$p = \frac{h}{\|p\|} p.$$
The matrix equation is solved by use of the Linpack subroutines DCHDC and DPOSL [39]. First DCHDC does a Cholesky decomposition on the matrix and then DPOSL solves the factored system.
4. Calculate the cut back factor α for the step. If $\|\alpha p\| < h$, set $h = 0.5h$, and $S^c = I$, and return to 1.
5. Calculate $x^+ = (x^c - \alpha p)^\wedge$. Where x^+ indicates the next step in the sequence of points defined by the optimization algorithm. If $f(x^+) \le f(x^c) - 10^{-4}\alpha \nabla f^T p$, then set $x^c = x^+$ and return to 1. Otherwise return to 4.

The line search algorithm [40] is as follows:

If no previous cutbacks for this p have been made, set $\alpha = 1$. If $x_i^+ = u_i$ and $x_i^c \ne u_i$ or $x_i^+ = l_i$ and $x_i^c \ne l_i$ for any i set $\alpha = 0.5\hat{\alpha}$. Where $\hat{\alpha}$ is the previous cut back factor that was used to calculate x^+. If $\hat{\alpha}$ is the first cut back factor for this p such that $x_i^+ \ne u_i$ if $x_i^c \ne u_i$ and $x_i^+ \ne l_i$ if $x_i^c \ne l_i$ for every i, then calculate β the unique minimizer of

$$(f(x^c - \hat{\alpha}p) - f(x^c) - \hat{\alpha}\nabla f(x^c)^T p)\beta^2 + \hat{\alpha}\nabla f(x^c)^T p\beta + f(x^c).$$

If $\beta \in [0.1, 0.5]$ set $\alpha = \beta\hat{\alpha}$. If $\beta > 0.5$ set $\alpha = 0.5\hat{\alpha}$. If $\beta < 0.1$ set $\alpha = 0.1\hat{\alpha}$.

Otherwise calculate γ the local minimizer to

$$a\gamma^3 + b\gamma^2 + \nabla f(x^c)^T p\gamma + f(x^c),$$

where:

$$\begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{\hat{\alpha} - \hat{\hat{\alpha}}} \begin{bmatrix} \frac{1}{\hat{\alpha}^2} & \frac{-1}{\hat{\hat{\alpha}}^2} \\ \frac{-\hat{\hat{\alpha}}}{\hat{\alpha}^2} & \frac{\hat{\alpha}}{\hat{\hat{\alpha}}^2} \end{bmatrix} \cdot \begin{bmatrix} f(x^c - \hat{\alpha}p) - f(x^c) - \nabla f(x^c)^T p\hat{\alpha} \\ f(x^c - \hat{\hat{\alpha}}p) - f(x^c) - \nabla f(x^c)^T p\hat{\hat{\alpha}} \end{bmatrix}.$$

Where $\hat{\hat{\alpha}}$ is the cutback factor previous to $\hat{\alpha}$. The local minimizer to this cubic polynomial is,

$$\gamma = \frac{-b + \sqrt{b^2 - 3a\nabla f(x^c)^T p}}{3a}$$

If $\gamma > 0.5\hat{\alpha}$ set $\alpha = 0.5\hat{\alpha}$. If $\gamma < 0.1\hat{\alpha}$ set $\alpha = 0.1\hat{\alpha}$. Otherwise $\alpha = \gamma$.

III. Yield Optimization Experiments

The simulator is used to conduct a series of yield optimization experiments. Three types of MESFET devices are selected: an ion implanted device, a buried channel device, and a uniform channel device. The ion implanted MESFET design is based upon a commercial device that had been empirically optimized by a standard procedure for maximum power-added efficiency. The device is capable of good RF performance and, as will be shown, the simulator is not able to alter the design to obtain significant maximum PAE improvements. The buried channel MESFET design is based upon an industrial device, but it is a nonoptimum prototype. The simulator significantly modifies the original design to improve performance. The starting design for the uniform channel MESFET is determined using standard, first principle design techniques. The initial design proves not to be optimum and is also significantly modified by the simulator to obtain improved performance. Starting from these MESFET designs, the yield of each device structure relative to to small-signal transducer gain, output power at 1 dB gain compression, and maximum power added efficiency is optimized. All devices are embedded in a 50 Ω circuit so that performance variations based upon device design, rather than circuit tuning conditions, can be investigated.

The experiments are conducted by selecting the experiment variables and estimating the disturbance covariance matrices. The variables are selected by performing a sensitivity analysis on each device structure. The variables to which the performance measures are most sensitive are included. These variables, for all the device types, are those that specify the gate geometry, the channel donor distribution, and the dc bias. The variables for each device type are listed in Table I. For these experiments the gate length is perturbed about the nominal manufacturable value which is held fixed and is not subject to optimization. Previous experience indicated that the optimizer always drives gate length to its minimum permitted value.

Each device's variables are assumed to be statistically independent. The covariance matrices are, therefore, diagonal. The diagonal elements are the squares of estimated standard deviations for each variable. In all cases, except the gate width, the standard deviations are taken to be 3% to 10% of the nominal values for initial designs. The gate width standard deviation is estimated to be the gate length standard deviation times the number of gate fingers. The variances are listed in Table I.

A number of other conditions are specified for the experiments. Three harmonics are used during harmonic balance calculations. The simulated circuit is reduced as shown in Fig. 3. The impedances presented to the gate and drain of the MESFET, as previously indicated, are 50 + $j0$ Ω at all harmonics. During each yield optimization, the number of sample devices for each yield estimate is 100. The number of sample devices in the presented yield histograms is 500.

In Figures 4, 5, and 6 the initial and optimized yield histograms are shown for ion implanted, buried channel, and uniform channel devices when small-signal transducer gain is the acceptance criterion. The optimizations require 151, 116, and 66 yield estimates, respectively. The three initial designs all have similar initial gain distributions with means varying from 5.1 to 5.9 (7.1 dB to 7.7 dB) while the standard deviations range from 0.7 to 1.1. The buried channel device exhibits the best improvement with a distribution mean increasing to 16.0. All the optimized distributions exhibit greater spread than their corresponding initial distributions. The standard deviations of the optimized designs range from 1.4 to 2.3. The increased spread results from the optimized designs being in a region where the gain is more sensitive to perturbations in the design parameters. For all three devices the optimizer changes the gate width the most. These changes improve the matching between the transistor and the 50 Ω circuit. This large change in gate width would not necessarily be expected for a tuned circuit. The buried and uniform channel devices exhibit large changes in their biases. For both devices, V_{GG} and V_{DD} increase. This bias shift is to a region of higher transconductance. The increased transconductance improves the small-signal gain.

Figs. 7, 8, and 9 show the initial and optimized yield histograms for ion implanted, buried channel, and uniform channel devices when output power at 1 dB gain compression is used as the acceptance criterion. The optimizations require 181, 180, and 152 yield estimates, respectively. The initial designs, again, have similar performance. The initial design distribution means range from 253 to 316 mW while the standard deviations vary from 55 to 71 mW. The uniform channel device shows the greatest increase in distribution mean to 2283 mW. However, the distribution spread is more than twice that of the two other optimized designs. In all three cases, the optimized designs show marked improvement over the initial designs. Again, the optimizer changes the gate widths of all three devices the most. However in this case, the increased gate widths allow more RF current to flow through the devices, thereby increasing the RF power. The impedance matching becomes a secondary consideration. For the buried and uniform devices V_{GG} and V_{DD} increase considerably. V_{DD} is also markedly increased for the ion implanted device. The biases change, along with the channel doping and gate width, so as to maximize the intersection of the 50 Ω load line and the device I-V curves.

Yield optimization using maximum power added efficiency as the acceptance criterion is also performed. Figs. 10, 11, and 12 show the initial and optimized yield histograms for ion implanted, buried channel, and uniform

TABLE I
DIAGONAL ELEMENTS IN DISTURBANCE COVARIANCE MATRICES FOR ION IMPLANTED DEVICE, BURIED CHANNEL DEVICE, AND UNIFORM DEVICE

Ion Implanted Device		Buried Channel Device		Uniform Channel Device	
Variable	Variance	Variable	Variance	Variable	Variance
L_g	1.6×10^{-3} $(\mu m)^2$	L_g	1.6×10^{-3} $(\mu m)^2$	L_g	1.6×10^{-3} $(\mu m)^2$
W_g	4.0×10^{-2} $(\mu m)^2$	W_g	2.6×10^{-2} $(\mu m)^2$	W_g	4.0×10^{-2} $(\mu m)^2$
E	5.6×10^{1} $(keV)^2$	t_{low}	1.7×10^{-6} $(\mu m)^2$	t_{chnl}	2.3×10^{-4} $(\mu m)^2$
D	1.4×10^{22} $(ions/cm^2)^2$	t_{high}	5.3×10^{-6} $(\mu m)^2$	N_{chnl}	2.5×10^{31} $(ions/cm^3)^2$
V_{GG}	1.0×10^{-2} $(V)^2$	N_{low}	5.6×10^{31} $(ions/cm^3)^2$	V_{GG}	2.5×10^{-3} $(V)^2$
V_{DD}	9.0×10^{-2} $(V)^2$	N_{high}	1.6×10^{33} $(ions/cm^3)^2$	V_{DD}	9.0×10^{-2} $(V)^2$
		V_{GG}	2.5×10^{-3} $(V)^2$		
		V_{DD}	9.0×10^{-2} $(V)^2$		

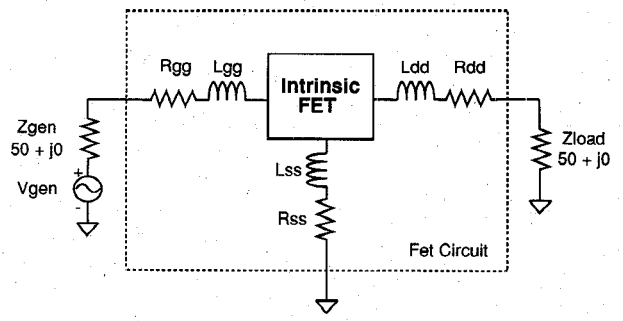
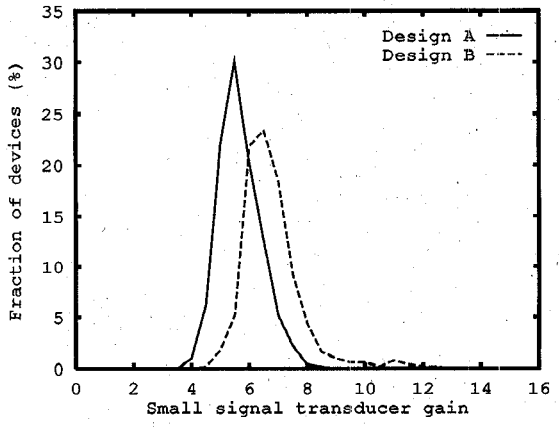


Fig. 3. The circuit simulated during yield optimizations. Referring to the full TEFLON schematic, the admittance blocks Yin and Yout are replaced with through networks, the impedance block Zs is replaced with a short circuit, and the Zfb block is replaced by an open circuit.

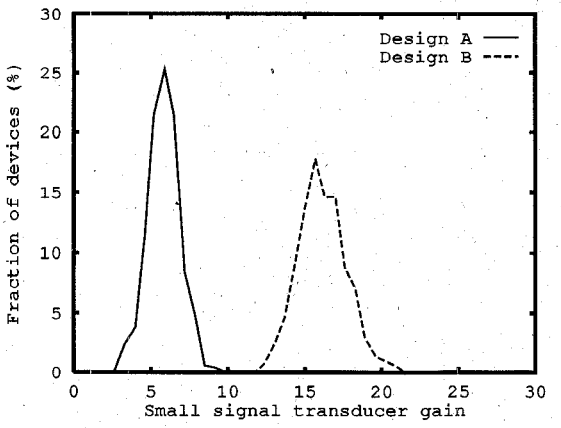


(a)

Parameter	Design	
	A	B
L_g (μm)	0.40	0.40
W_g (μm)	1000	609
E (keV)	150	133
D $\left(\frac{ions}{cm^2}\right)$	4.0×10^{12}	3.7×10^{12}
V_{GG} (V)	-2.0	-2.1
V_{DD} (V)	6.0	6.2

(b)

Fig. 4. Results of yield optimization on a ion implanted device using small-signal transducer gain criteria. (a) Simulated histograms: Design A—initial design. Design B—optimized design. (b) Initial and optimized device parameter values.



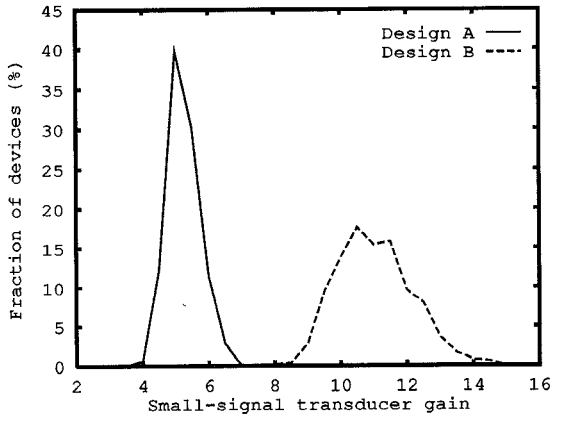
(a)

Parameter	Design	
	A	B
L_g (μm)	0.40	0.40
W_g (μm)	800	401
t_{low} (μm)	0.026	0.026
t_{high} (μm)	0.046	0.062
N_{low} $\left(\frac{ions}{cm^3}\right)$	1.5×10^{17}	1.2×10^{17}
N_{high} $\left(\frac{ions}{cm^3}\right)$	8.0×10^{17}	7.9×10^{17}
V_{GG} (V)	-1.0	-0.13
V_{DD} (V)	6.0	7.5

(b)

Fig. 5. Results of yield optimization on a buried channel device using small-signal transducer gain criteria. (a) Simulated histograms: Design A—initial design. Design B—optimized design. (b) Initial and optimized device parameter values.

channel devices. The optimizations require 143, 63, and 203 yield estimates, respectively. The initial designs in this case are not all similar. The initial uniform channel design has a distribution mean of 19%, whereas, the initial ion implanted and buried channel designs have distribution means of 38.5% and 40.0%, respectively. Yield optimization improved the uniform channel design's distribution mean to 40.9%, but the optimized buried channel design has the best performance with a distribution mean of 47.6%. No improvement is noted in the ion im-

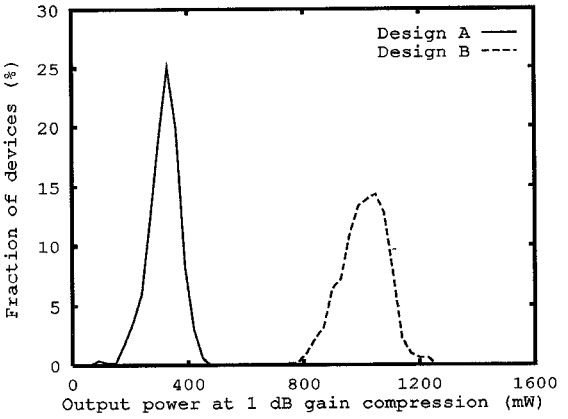


(a)

Parameter	Design	
	A	B
L_g (μm)	0.40	0.40
W_g (μm)	1000	400
t_{chnl} (μm)	0.30	0.22
N_{chnl} $\left(\frac{ions}{cm^3}\right)$	1.0×10^{17}	9.0×10^{16}
V_{GG} (V)	-1.0	-0.29
V_{DD} (V)	6.0	12.3

(b)

Fig. 6. Results of yield optimization on a uniform channel device using small-signal transducer gain criteria. (a) Simulated histograms: Design A—initial design. Design B—optimized design. (b) Initial and optimized device parameter values.

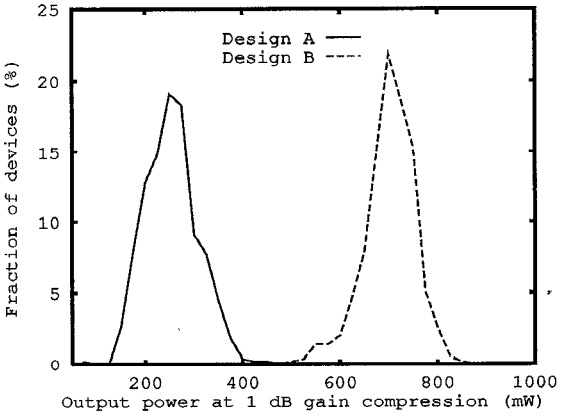


(a)

Parameter	Design	
	A	B
L_g (μm)	0.40	0.40
W_g (μm)	1000	2755
E (keV)	150	171
D $\left(\frac{ions}{cm^2}\right)$	4.0×10^{12}	3.3×10^{12}
V_{GG} (V)	-2.0	-2.0
V_{DD} (V)	6.0	9.6

(b)

Fig. 7. Results of yield optimization on a ion implanted device using 1 dB gain compression output power criteria. (a) Simulated histograms: Design A—initial design. Design B—optimized design. (b) Initial and optimized device parameter values.



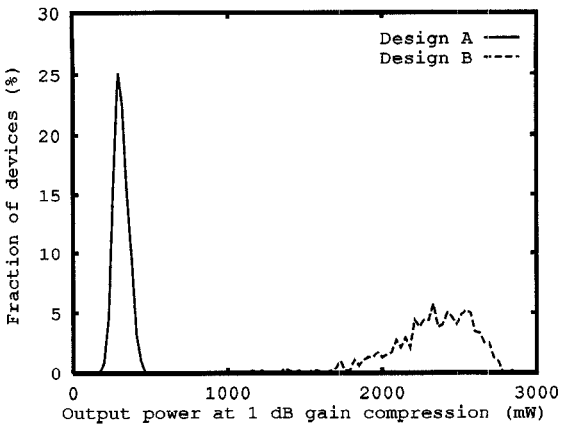
(a)

Parameter	Design	
	A	B
L_g (μm)	0.40	0.40
W_g (μm)	800	2298
t_{low} (μm)	0.026	0.026
t_{high} (μm)	0.046	0.049
N_{low} $\left(\frac{ions}{cm^3}\right)$	1.5×10^{17}	1.4×10^{17}
N_{high} $\left(\frac{ions}{cm^3}\right)$	8.0×10^{17}	7.8×10^{17}
V_{GG} (V)	-1.0	-0.65
V_{DD} (V)	6.0	10.0

(b)

Fig. 8. Results of yield optimization on a buried channel device using 1 dB gain compression output power criteria. (a) Simulated histograms: Design A—initial design. Design B—optimized design. (b) Initial and optimized device parameter values.

planted design. This result is anticipated since the ion implanted initial design is based on a mature industry device which has been empirically optimized for maximum power added efficiency. As is the case with the gain and output power at 1 dB gain compression, the largest changes occur in the gate width and the dc bias. The changes in gate width improve the match to the 50 Ω circuit. The bias points shift in such a way as to minimize the dc power supplied to the device while maximizing the intersection of the 50 Ω load line and the device I-V curves.

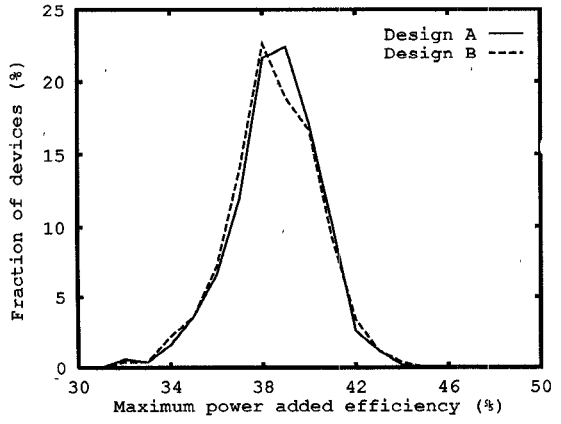


(a)

Parameter	Design	
	A	B
L_g (μm)	0.40	0.40
W_g (μm)	1000	2952
t_{chnl} (μm)	0.30	0.22
N_{chnl} $\left(\frac{ions}{cm^3}\right)$	1.0×10^{17}	7.3×10^{16}
V_{GG} (V)	-1.0	-0.44
V_{DD} (V)	6.0	16.0

(b)

Fig. 9. Results of yield optimization on a uniform channel device using 1 dB gain compression output power criteria. (a) Simulated histograms: Design A—initial design. Design B—optimized design. (b) Initial and optimized device parameter values.

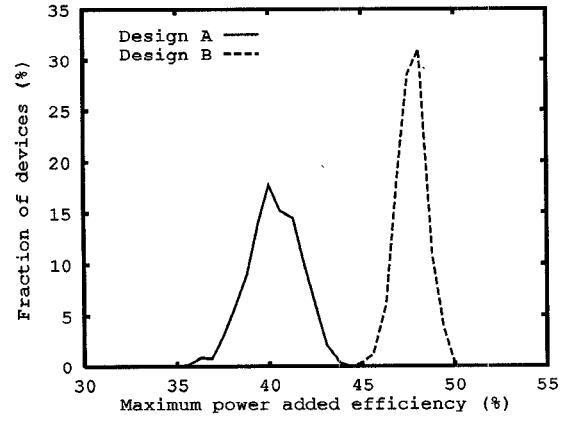


(a)

Parameter	Design	
	A	B
L_g (μm)	0.40	0.40
W_g (μm)	1000	814
E (keV)	150	159
D $\left(\frac{ions}{cm^2}\right)$	4.0×10^{12}	4.1×10^{12}
V_{GG} (V)	-2.0	-2.2
V_{DD} (V)	6.0	5.7

(b)

Fig. 10. Results of yield optimization on a ion implanted device using maximum power added efficiency criteria. (a) Simulated histograms: Design A—initial design. Design B—optimized design. (b) Initial and optimized device parameter values.



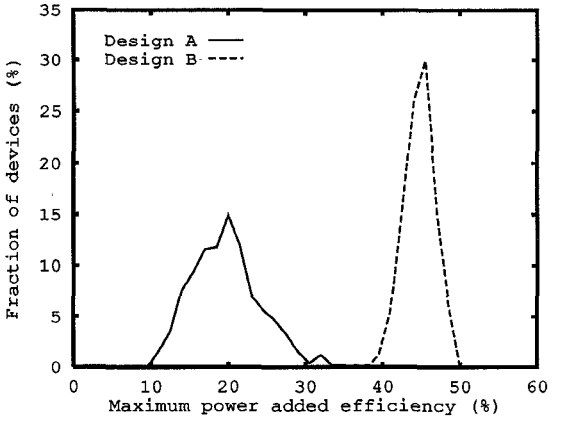
(a)

Parameter	Design	
	A	B
L_g (μm)	0.40	0.40
W_g (μm)	800	380
t_{low} (μm)	0.026	0.027
t_{high} (μm)	0.046	0.044
N_{low} $\left(\frac{ions}{cm^3}\right)$	1.5×10^{17}	1.3×10^{17}
N_{high} $\left(\frac{ions}{cm^3}\right)$	8.0×10^{17}	7.8×10^{17}
V_{GG} (V)	-1.0	-0.43
V_{DD} (V)	6.0	5.8

(b)

Fig. 11. Results of yield optimization on a buried channel device using maximum power added efficiency criteria. (a) Simulated histograms: Design A—initial design. Design B—optimized design. (b) Initial and optimized device parameter values.

Different optimum device designs result for each of the specified performance criteria. That is, an optimum PAE design is different from an optimum design for either maximum output power or gain. This series of experiments indicates that the buried channel device is the best device structure of the three when gain and power added efficiency are of primary concern. The uniform channel device gives the best average performance for output power at 1 dB gain compression, but the performance distribution exhibits excessive variance. These result, of



(a)

Parameter	Design	
	A	B
L_g (μm)	0.40	0.40
W_g (μm)	1000	968
t_{chnl} (μm)	0.30	0.25
N_{chnl} $\left(\frac{ions}{cm^3}\right)$	1.0×10^{17}	9.0×10^{16}
V_{GG} (V)	-1.0	-1.8
V_{DD} (V)	6.0	6.9

(b)

Fig. 12. Results of yield optimization on a uniform channel device using maximum power added efficiency criteria. (a) Simulated histograms: Design A—initial design. Design B—optimized design. (b) Initial and optimized device parameter values.

course, only apply to the devices embedded in a 50 Ω circuit. Different results are possible when circuit tuning conditions are considered.

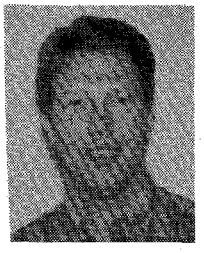
IV. Conclusion

A large-signal physics based GaAs MESFET and circuit simulator for yield estimation and optimization has been developed. The simulator's yield pass-fail criteria are small- and large-signal performance measures for power amplifiers. For investigation of ion-implanted devices SUPREM 3.5, a GaAs process simulator, has been incorporated so that pertinent process variables directly specify the MESFET's channel donor distribution. The yield is estimated with a Monte Carlo algorithm. The optimizer used for maximizing the yield estimate is a quasi-Newton algorithm which uses decreasing gradient scales to overcome the estimate's Monte Carlo noise. The optimizer can alter the device design, RF circuit, and operating parameters until a maximized yield for a given performance specification is achieved. In this manner design centering can be performed. Since the device model is physics based, design optimization can be performed before fabrication. Use of the simulator should allow significant reductions in the time and costs required to produce MIMIC circuits.

References

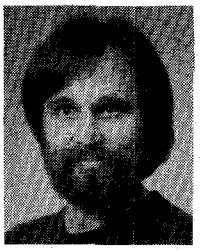
[1] T. R. Scott and T. P. Walker, "Regionalization: a method for generating joint density estimates," *IEEE Trans. Circuits Syst.*, vol. 23, pp. 229–234, Apr. 1976.
[2] K. H. Leung and R. Spence, "Efficient statistical circuit analysis," *Electron. Lett.*, vol. 10, pp. 360–362, 1974.
[3] S. W. Director and G. D. Hachtel, "The simplicial approximation approach to design centering," *IEEE Trans. Circuits Syst.*, vol. 24, pp. 363–372, 1977.
[4] ——, "Computationally efficient yield estimation procedures based on simplicial approximation," *IEEE Trans. Circuits Syst.*, vol. 25, pp. 121–130, 1978.
[5] S. W. Director, W. Maley, and A. J. Strojwas, *VLSI Design for Manufacturing: Yield Enhancement*. Boston: Kluwer, 1988.
[6] ——, "Yield optimization for arbitrary statistical distributions: Part I—Theory," *IEEE Trans. Circuits Syst.*, vol. 27, pp. 245–253, Apr. 1980.
[7] H. L. Abdel-Malek and J. W. Bandler, "Yield optimization for arbitrary statistical distributions: Part II—Implementation," *IEEE Trans. Circuits Syst.*, vol. 27, pp. 253–262, Apr. 1980.
[8] J. W. Bandler and H. L. Abdel-Malek, "Optimal centering, tolerancing, and yield determination via updated approximations and cuts," *IEEE Trans. Circuits Syst.*, vol. 25, pp. 853–871, Oct. 1978.
[9] H. L. Abdel-Malek and A. S. O. Hassan, "The ellipsoidal technique for design centering and region approximation," *IEEE Trans. Computer-Aided Design*, vol. 10, pp. 1006–1014, Aug. 1991.
[10] R. S. Soin and R. Spence, "Statistical exploration approach to design centering," *Proc. Inst. Elec. Eng.*, vol. 127, pt. G., pp. 260–269, 1980.
[11] M. A. Styblinski and A. Ruszczynski, "Stochastic approximation approach to statistical circuit design," *Electron. Lett.*, vol. 19, pp. 300–302, 1983.
[12] G. Kjellström and L. Taxèn, "Stochastic optimization in system design," *IEEE Trans. Circuits Syst.*, vol. 28, pp. 702–715, July 1981.
[13] K. Singhal and J. F. Pinel, "Statistical design centering and tolerancing using parametric sampling," *IEEE Trans. Circuits Syst.*, vol. 28, pp. 692–702, July 1981.
[14] J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods*. London: Methuen, 1964.
[15] D. E. Hocevar, M. R. Lightner, and T. N. Trick, "A study of variance reduction techniques for estimating circuit yields," *IEEE Trans. Computer-Aided Design*, vol. 2, pp. 180–192, July 1983.
[16] R. S. Soin and P. J. Rankin, "Efficient tolerance analysis using control variates," in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 1, pp. 143–148, 1981.
[17] P. J. Rankin and R. S. Soin, "Efficient Monte Carlo yield prediction using control variates," *Inst. Elec. Eng. Proc. Part G*, vol. 132, pp. 131–142, 1985.
[18] R. M. Biernacki and J. W. Bandler, "Efficient quadratic approximation for statistical design," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 1449–1454, Nov. 1989.
[19] J. W. Bandler and S. H. Chen, "Circuit optimization: the state of the art," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 424–443, Feb. 1988.
[20] J. W. Bandler, Q-J. Zhang, J. Song, and R. M. Biernacki, "FAST gradient based yield optimization of nonlinear circuits," *IEEE Trans. Microwave Theory Tech.*, vol. 38, pp. 1701–1710, Nov. 1990.
[21] J. W. Bandler, Q. Cai, R. M. Biernacki, S. H. Chen, and Q-J. Zhang, "Physics-based design and yield optimization of MMICs," in *Proc. 21st European Microwave Conf.*, pp. 1515–1520, 1991.
[22] M. A. Khatibzadeh and R. J. Trew, "A large-signal analytical model for the GaAs MESFET," *IEEE Trans. Microwave Theory and Techniques*, vol. 36, pp. 231–238, 1988.
[23] M. A. Khatibzadeh, "Large-signal modeling of gallium arsenide field-effect transistors," Ph.D. dissertation, North Carolina State University, Raleigh, 1987.
[24] A. J. Strojwas, Ed., *Statistical Design of Integrated Circuits*. New York: IEEE Press, 1987.
[25] R. Spence and R. Soin, *Tolerance Design of Electronic Circuits*. Reading, MA: Addison-Wesley, 1988.
[26] R. Brayton and R. Spence, *Sensitivity and Optimization*. New York: Elsevier, 1980.
[27] P. W. Becker and F. Jensen, *Design of Systems and Circuits of Maximum Reliability of Maximum Production Yield*. New York: McGraw-Hill, 1977.
[28] J. Purviance and M. Meehan, "CAD for Statistical Analysis and De-

sign of Microwave Circuits," *Int. J. Microwave and Millimeter-Wave Computer-Aided Engineering*, vol. 1, pp. 59–76, 1991.
[29] M. D. Deal, S. E. Hansen, R. Anholt, S. Chou, J. D. Plummer, P. W. Dutton, T. W. Sigmon, D. A. Stevenson, C. R. Helms, and J. C. Bravman, "Suprem 3.5, process modeling of gallium arsenide," in *1987 Int. Electron Devices Meeting Technical Dig.*, 1987, pp. 256–259.
[30] W. R. Frensley, "Power-limiting breakdown effects in GaAs MESFETs," *IEEE Trans. Electron Dev.*, vol. 28, pp. 962–970, Aug. 1981.
[31] H. Q. Tserng, W. R. Frensley, and P. Sauner, "Light emission of GaAs power MESFETs under RF drive," *IEEE Electron Dev. Lett.*, vol. 1, pp. 20–21, Feb. 1980.
[32] T. A. Winslow, P. Gilmore, R. J. Trew, and C. T. Kelley, "Doping profiles for optimum class B performance of GaAs MESFET amplifiers," in *Proc. IEEE/Cornell Confer. Advanced Concepts in High Speed Semiconductor Devices and Circuits*, 1991.
[33] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*. Cambridge: Cambridge University Press, 1986.
[34] J. Purviance, D. Criss, and D. Monteith, "FET model statistics and their effects on design centering and yield prediction for microwave amplifiers," in *1988 IEEE MTT-S Int. Microwave Symp. Dig.*, pp. 315–318.
[35] G. L. Bilbro, M. B. Steer, R. J. Trew, C.-R. Chang, and S. G. Skaggs, "Extraction of the parameters of equivalent circuits of microwave transistors using tree annealing," *IEEE Trans. Microwave Theory Techniques*, vol. 38, pp. 1711–1718, Nov. 1990.
[36] R. J. Gilmore, M. Eron, and T. Zhang, "Yield optimization of a MMIC distributed amplifier using physically-based device models," in *1991 IEEE MTT-S Int. Microwave Symp. Dig.*, 1991, pp. 1205–1208.
[37] M. D. Meehan, "Statistical techniques for objective characterization of microwave device statistical data," in *1991 IEEE MTT-S Int. Microwave Symp. Dig.*, pp. 1209–1212.
[38] W. Feller, *An Introduction to Probability Theory and its Applications, Vol. 2*. New York: Wiley, 1966.
[39] J. J. Dongarra, C. B. Moler, J. R. Bunch, and G. W. Stewart, *Linpack Users' Guide*. Philadelphia: SIAM, 1979.
[40] J. E. Dennis, Jr. and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Englewood Cliffs, NJ: Prentice-Hall, 1983.



Dan E. Stoneking (S'80-M'81) received the Bachelors, Masters, and Ph.D. degrees in electrical engineering from North Carolina State University, Raleigh, in 1981, 1986, and 1992, respectively.

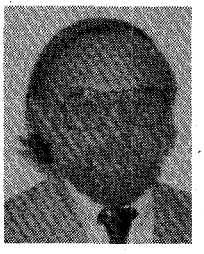
From 1981 to 1983 he was with United Technologies Mostek in Carrollton, TX. While with Mostek, he designed CMOS switched capacitor filters for telephony applications. From 1983 to 1984 he worked at the General Electric Electronics Laboratory in Syracuse, NY where he designed CMOS logic circuits for a serial arithmetic processor/controller chip used in phased array radar systems. In 1992 he joined M/A-COM Corporation in Lowell, MA to support nonlinear modeling and simulation efforts. His research interests include design, modeling, and optimization of nonlinear microwave circuits, statistical modeling, and expert systems for microwave circuit synthesis and simulation.



Griff L. Bilbro received the B.S. degree in 1973 from Case Western Reserve University in Cleveland, OH, and the Ph.D. degree in 1977 from the University of Illinois at Urbana-Champaign, where he was a National Science Foundation Graduate Fellow in physics.

Until 1984, he designed computer models of complex systems in industry. He joined the Department of Electrical and Computer Engineering at North Carolina State University where he is now a Visiting Assistant Professor. He is a member of the Center for Communications and Signal Processing as well as the High Frequency Laboratory at NCSU where he develops practical engineering techniques for global optimization in modeling, device physics, design centering, and image processing.

Paul A. Gilmore received his Bachelors degree in mathematics from Pacific Lutheran University in 1984 and his Masters degree in mathematics from Washington St. University in 1986. He is currently working on his Ph.D. in applied mathematics in the area of optimization under the direction of Dr. C.T. Kelley at North Carolina State University.



Robert J. Trew (S'71-M'74-SM'87-F'91) received the Ph.D. degree in electrical engineering from the University of Michigan in 1975.

From 1975 to 1977 he was employed by the Watkins-Johnson Company in Palo Alto, CA where he was involved in the design and development of various microwave solid-state components. In 1977 he joined the facity of North Carolina State University, Raleigh where he is currently Professor of Electrical and Computer Engineering and Director of the High Frequency Electronics Laboratory. His research interests are in the areas of microwave/millimeter wave solid-state device modeling, microwave device and circuit design, high frequency experimental characterization, and parameter extraction techniques.

Dr. Trew is a member of Eta Kappa Nu, Sigma Xi, Tau Beta Pi, the Electromagnetics Academy, and the American Association for the Advancement of Science.

C. T. Kelley received the B.A. degree in mathematics from Vanderbilt University in 1973 and the Ph.D. degree in applied mathematics from Purdue University in 1976.

In the 1977–1978 academic year he was an Assistant Scientist at the Army Mathematics Research Center of the University of Wisconsin. He joined the Mathematics department of North Carolina State University in 1978 where is currently Professor. His current research interests include numerical methods for optimization, multilevel methods for integral and differential equations, and optimal control.

Dr. Kelley is a member of SIAM, the Mathematical Programming Society, the IEEE Computer Society, and ACM/SIGNUM.