Parameter Extraction of Microwave Transistors using Tree Annealing

by

Michael B. Steer, Griff L. Bilbro, Robert J. Trew, and Steven G. Skaggs

High Frequency Electronics Laboratory and
the Center for Communications and Signal Processing
Department of Electrical and Computer Engineering, North Carolina State University,
Raleigh, NC 27695-7911.

INTRODUCTION

As quantitative accuracy in the design phase becomes critical, it becomes increasingly important to develop accurate equivalent circuit models based upon experimental results and physical insight. Although the equivalent circuit characterization technique is widely used, the lack of a general purpose global optimization scheme has prevented the application of realistic, physically based equivalent circuits and the incorporation of physical insight into the optimization process.

Gradient descent algorithms are most commonly used in parameter extraction, but even when these are combined with random search techniques results are seldom satisfactory. An alternative technique has recently been proposed by Vai et al [1] who use stochastic simulated annealing. This is a near global optimization technique but known to be unsuitable for many continuous optimization problems. We have developed an alternative formulation of simulated annealing using a tree-based Metropolis procedure, so called tree annealing. Tree annealing is suited to continuous optimization problems and in particular to transistor parameter extraction.

SIMULATED ANNEALING IN COMBINATORIC AND CONTINUOUS OPTIMIZATION

Simulated Annealing (SA) is a smart random search technique which is more efficient than exhaustive search yet more robust than gradient descent. By gradually decreasing a control parameter (usually called temperature T), the search is systematically concentrated into regions likely to contain a global minimum, but is still random enough to escape most local minima. In some cases global convergence is provable [4]. Instead of directly minimizing an objective function $E(x)$, SA samples an associated probability density $p_T(x) \approx \exp(-E(x)/T)$ at a low enough value of T to make $p_T(x)$ sharply peaked at the x that minimizes $E(x)$. Kirkpatrick et al [6] observed that the Metropolis procedure [8] could do this sampling efficiently enough to be practical if it was started at high T then gradually "cooled" to the final T.

SA is widely used in combinatoric optimization but less commonly in continuous optimization [7]. However, Vai et. al. have recently applied SA to the continuous optimization problem of estimating the parameter values of a 7 or 8 element equivalent circuit for a MODFET [1]. In SA a current state x is randomly perturbed into a candidate state x'.

CH2790-4/89/0000-0195 $1.00 © 1989 IEEE

The candidate is either accepted or rejected by the Metropolis criterion, that is with probability $\min(1, p_T(x')/p_T(x))$. If the candidate is accepted, it replaces the current state in the next time step; otherwise the current state persists and the candidate is discarded. The resulting sequence of states is guaranteed [8] to asymptotically sample the density p_T.

Vanderbilt and Louie [5] have reported SA generally becomes inefficient for continuous optimization as the number of unknowns increases. In their treatment of continuous optimization problems, they extended classic simulated annealing to 10 unknowns. Moreover if $E(x)$ is "too irregular" [2], SA chooses between deep but widely separated local minima before it can determine which is deeper.

TREE ANNEALING THEORY

We have solved the inefficiency and irregularity problems with a reformulation of SA that can robustly extract a 14 parameter transistor equivalent circuit from data. Our reformulation generates candidates from a global approximation of $p_T(x)$ rather than by perturbation of the current state. We call this global approximation $g_T(x)$ because it is used to generate candidates. We maintain g_T in a k-d binary tree [3] data structure which also allows us to conveniently sample from g_T. Because of this tree, we call our algorithm Tree Annealing (TA). Such a tree-based Metropolis Procedure is efficient and is able to sample irregular densities, but it requires an unfamiliar form of the Metropolis Criterion [9]. Like the familiar Metropolis procedure, ours can be shown to asymptotically sample $p_T(x)$. Because of this guarantee holds even when $g_T(x)$ is incorrect (though efficiency is reduced by a poor g_T), we can accumulate the sampled x's to improve an incorrect g_T. Furthermore, in the final stages of the algorithm, these samples can be used to estimate the precision with which the optimal values of x have been determined.

We have run TA under a variety of cooling schedules and problems. Except for the tree-based Metropolis Procedure, TA behaves qualitatively like SA. Our tree is built up of nodes containing four fields: n_l, n_r, the number of previous samples from the left and right subtrees, and l_c, l_r, pointers to the left and right subtrees. Whenever we create a new node (a leaf), we set $n_r = n_l = 1$ and we set $c_l = c_r =$ NULL. Implicitly associated with each node is a portion of the parameter space to be searched; with the root node associated with the entire parameter space as specified by the user. The space associated with a node is divided equally between its two children in a predetermined way. A node with no children ($c_l = c_r =$ NULL) is a leaf and other nodes are interior nodes.

Sampling from $g_T(x)$ always begins at the root node and terminates at a leaf. At an interior node, a random decision is made to choose the right child with probability $n_r/(n_r + n_l)$ and otherwise to descend to the left child. At a leaf node, a random candidate vector x' is drawn uniformly from the portion of parameter space associated with that leaf. This candidate is compared with the current x and one of them is selected to be the next sample x'' according to our modified Metropolis Criterion: the probability of accepting x', given the current sample x, is $\min(1, g_T(x) * p_T(x')/(g_T(x') * p_T(x)))$. Whether or not x' is accepted, both its children are created at c_l and c_r; this causes the tree to grow into regions of higher probability and permits subsequent sampling to concentrate there. Furthermore every node in the path from the root to the leaf containing x'' is then modified by incrementing its n_r if the right child was selected on the way to x'', otherwise by incrementing n_l. This completes the generation of one sample from g_T and

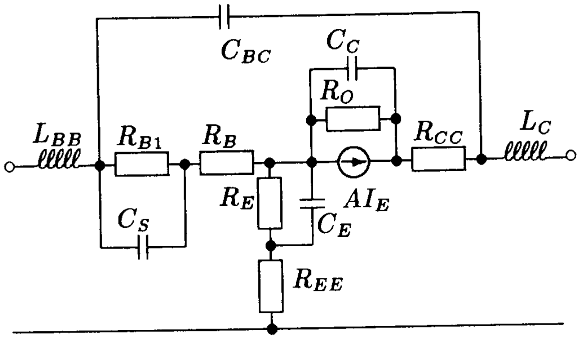

Figure 1: Equivalent circuit of a HBT.

the modification of g_T to record that sample.

RESULTS

The tree annealing optimization algorithm was used to extract the parameters of the HBT transistor of Mishra et al [10] using the physically based equivalent circuit of Fig. 1 and deembedded scattering (S) parameter measurements from 45 MHz to 26.5 GHz. In the transistor model the current gain

$$A = \frac{A(0)}{1 + \mathrm{j}f/F} e^{-\mathrm{j}2\pi f \tau_C'}$$

where $A(0)$ is the DC current gain, τ_C' is the carrier transit time through the base, f is frequency and F accounts for the variation in the base width and is the frequency at which the magnitude of the internal current gain is 3 dB down. F can not be measured directly and is determined from the measured unity current gain frequency f_t and the time constants in the equivalent circuit [11]:

$$F = 1/(2\pi\tau_F)$$

where

$$\tau_F = \tau_t - (\tau_C' + R_E C_E + R_E C_E + R_C C_C)$$

and

$$\tau_t = 1/(2\pi f_t)$$

The extracted model element values are listed in table 1 and in Fig. 2 the measured stability factor, k, and maximum available gain, Gmax, are compared to those evaluated from the equivalent circuit. The agreement is excellent indicating an accurate parameter extraction.

PARAMETER	OPTIMIZED VALUE	STANDARD DEVIATION
L_{BB}	41.4 pH	3.2 pH
R_{B1}	50.2 Ω	4.1 Ω
C_S	57.5 fF	4.5 fF
R_B	18.0 Ω	1.1 Ω
R_E	2.11 Ω	0.24 Ω
C_E	71.9 fF	12.3 fF
R_{EE}	2.73 Ω	0.20 Ω
$A(0)$	0.973	¡0.005
τ'_C	1.07 ps	0.10 ps
C_C	24.7 fF	1.7 fF
R_O	30.7 kΩ	3.8 kΩ
R_C	3.70 Ω	0.54 Ω
L_C	63.2 pH	10.2 pH
C_{BC}	37.4 fF	1.5 fF

Table 1: Values and standard deviations of optimized elements. The standard deviations are for 10% fractional errors in the measured S parameters.

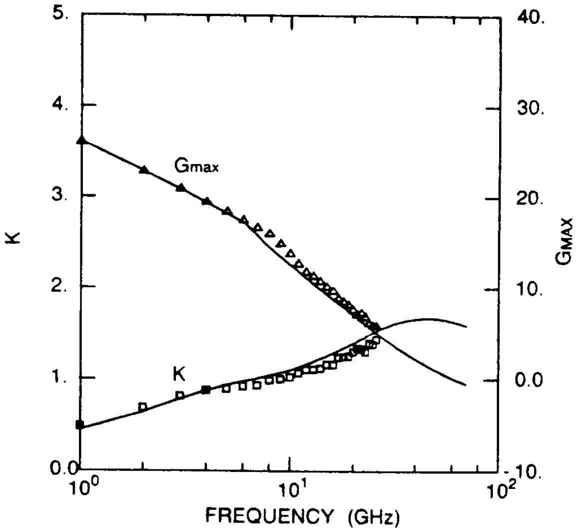

Figure 2: Comparison of calculated stability factor, k, and maximum available gain, Gmax, with those measured.

	L_{BB}	R_{B1}	C_S	R_B	R_E	C_E	R_{EE}	$A(0)$	τ'_C	C_C	R_O	R_C	L_C	C_{BC}
L_{BB}	1.00													
R_{B1}	-0.21	1.00												
C_S	-0.50	0.15	1.00											
R_B	-0.15	0.06	0.49	1.00										
R_E	0.26	-0.28	0.02	0.23	1.00									
C_E	0.35	0.07	-0.47	-0.24	-0.19	1.00								
R_{EE}	-0.24	-0.29	0.19	-0.15	-0.35	0.05	1.00							
$A(0)$	-0.07	0.14	0.12	0.02	0.08	0.04	0.12	1.00						
τ'_C	-0.28	0.01	-0.01	-0.23	0.03	-0.14	0.33	0.01	1.00					
C_C	-0.24	-0.24	0.59	0.49	0.10	-0.60	0.15	0.14	-0.00	1.00				
R_O	-0.17	-0.38	0.20	0.01	-0.11	-0.20	0.40	-0.07	0.05	0.30	1.00			
R_C	0.21	0.36	-0.16	0.03	-0.09	0.27	-0.41	-0.08	-0.32	0.27	-0.35	1.00		
L_C	0.10	0.26	-0.45	-0.36	-0.08	0.46	-0.14	-0.07	0.05	0.68	-0.32	0.13	1.00	
C_{BC}	0.04	0.24	-0.40	-0.35	-0.09	0.26	-0.08	-0.18	0.15	0.80	-0.19	0.20	0.24	1.00

Table 2: Correlations of calculated equivalent circuit elements for a 10% fractional errors in the measured S parameters.

Also given in table 1 are the standard deviations (σ) of the optimized element parameters for a 10% variation in the fractional error of the two-port admittances. Small σ implies that the error function is sensitive to variations in this element and so the element value is accurately determined to a tight tolerance. The provision of error estimates indicates the importance of an element on the objective function minimization and provides useful information to device and circuit designers. This is particularly important in the tolerancing of MMIC's as it directs effort towards the control of elements with small σ as they have most effect on external characteristics.

Various combinations of the elements of a transistor equivalent circuit are closely correlated as shown in table 2. In this table the cross correlation coefficients of the optimized elements are shown for 10% fractional errors in the measured S parameters. These statistics are immediately available in the tree annealing procedure. It can be seen that the dc current gain A(0) is precisely determined and, as expected, R_E and R_{EE} are tightly correlated. Also most of the capacitors are tightly correlated. A high coefficient indicates that the associated elements can not be adequately resolved. If it is important to distinguish them, additional or independent measurements are required.

CONCLUSION

Good results were obtained from the parameter extraction technique and the ability of MFA not to get locked in local minima enabled a physically based equivalent circuit model to be used. Tree annealing is essentially a smart random search technique and so requires many more functional evaluations than do gradient based minimization algorithms. However, no starting guess is required and the bounds on parameter values can be widely separated with little effect on optimization time.

ACKNOWLEDGEMENT

This work was supported in part by a National Science Foundation Presidential Young Investigator Award Grant No. ECS-8657836 to M.B. Steer.

REFERENCES

[1] M-K. Vai, S. Prasad, N.C. Li, and F. Kai, "Modeling of microwave semiconductor devices using simulated annealing optimization," *IEEE Trans. Electron Devices*, Vol. ED-36, April 1989, pp. 761-762.

[2] R.A. Rutenbar, "Simulated annealing algorithms: an overview," *IEEE Circuits and Devices Magazine*, Jan. 1989, pp. 19-26.

[3] S.M. Omohundro, "Efficient algorithms with neural network behaviour," Report No. UIUCDCS-R-87-1331, Dept. of Computer Science, University of Illinois at Urbana-Champaign, Urbana, Illinois, 1987.

[4] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Proc. Pattern Analysis and Machine Intel.*, Vol. PAMI-6, 1984, pp. 721-741.

[5] D. Vanderbilt and S. G. Louie, "A Monte Carlo simulated annealing approach to optimization over continuous variables," *J. Comput. Phys.*, Vol. 36, 1984, pp. 259-271.

[6] S. Kirkpatrick, C. D. Gelatt and M.P. Vecchi, "Optimization by Simulated Annealing," *Science*, Vol. 220, 1983, pp. 671-680.

[7] P.J.M. Van Laarhoven and E.H. Aarts, *Simulated Annealing: Theory and Applications*," Reidel, Dordrecht, Holland, 1987.

[8] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller and E. Teller, "Equation of state calculations by fast computing machines, *J. of Chem. Physics*, Vol. 21, 1953, pp. 1087-1092.

[9] M.H. Kalos and P. A. Whitlock, *Monte Carlo Methods, Volume 1: Basics*, Wiley & Sons, New York, 1986.

[10] U.K. Mishra, A.S. Brown, and S.E. Rosenbaum, "D.C. and R.F. Performance of 0.1μ Gate Length $Al_{.48}In_{.52}As - Ga_{.38}In_{.62}As$ Pseodomorphic HEMTs," *IEDM Technical Digest*, December 1988, pp. 180-183.

[11] R.J. Trew, U.K. Mishra, W.L. Pribble and J.F. Jensen, "A parameter extraction technique for heterojunction bipolar transistors," *IEEE MTT-S International Microwave Symposium Digest*, June 1989, pp. 897-900.