ROBUST MODEL PARAMETER EXTRACTION USING LARGE-SCALE OPTIMIZATION CONCEPTS

J.W. Bandler*, S.H. Chen, S. Ye* and Q.J. Zhang*

Optimization Systems Associates Inc.
163 Watson's Lane
Dundas, Ontario, Canada L9H 6L1

ABSTRACT

A robust approach to model parameter extraction is presented. This approach utilizes multi-bias measurements and dc device characteristics. Novel automatic decomposition concepts for large-scale optimization are used to detect possible model topology deficiencies. Powerful ℓ_1 optimization is employed with adjoint analyses for both dc and ac sensitivities.

INTRODUCTION

This paper describes a robust approach which substantially expands the multi-circuit algorithm introduced in [1-2]. It also exploits the automatic decomposition concepts for large-scale optimization proposed in [3].

Conventionally, we seek a set of model parameters which minimizes the difference between the model responses and the measurements. To alleviate indeterminacy as well as for simplicity, techniques have been implemented (e.g., [4-5]) which separate the dc, low- and high-frequency measurements and divide the model parameters into corresponding subsets. Such a sequential decoupling approach may not be reliable: parameters determined solely from dc measurements may not be suitable for the purpose of microwave simulation, and the information contained in rf measurements is not fully utilized.

The multi-circuit algorithm [1-2] can improve the uniqueness of the solution by simultaneously processing multiple sets of S-parameter measurements made under different bias conditions. However, model parameters were assumed to be either completely bias-independent or arbitrarily bias-dependent.

Our new approach not only attempts to match dc and ac measurements simultaneously, but also employs the dc characteristics of the device as constraints on the bias-dependent parameters. These constraints reduce the degrees of freedom in modeling, thus improving the uniqueness and reliability of the solution.

Bandler and Zhang [3] have proposed a decomposition dictionary to reveal the interdependency between model responses and model parameters. We exploit such a dictionary to examine a sequence of increasingly more complex models. We start with the simplest model, subsequently adding elements according to the dictionary for a better match between the model responses and the measurements.

* J.W. Bandler is also with and S. Ye and Q.J. Zhang are with the Simulation Optimization Systems Research Laboratory and the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada L8S 4L7.

ℓ_1 optimization is highly favored for device modeling [1]. We have integrated a powerful ℓ_1 algorithm [6] into our new approach. To provide gradients efficiently, adjoint analyses are performed to obtaining both dc and ac sensitivities.

MULTI-BIAS FORMULATION WITH DC CONSTRAINTS

Consider a device model with its equivalent circuit. The model parameters are classified as bias-independent, unconstrained bias-dependent, and constrained bias-dependent. We also separate the parameters that appear in both dc and ac (small-signal) models from those appearing only in the ac model. Therefore, we define six subsets of model parameters denoted by $\boldsymbol{\phi}_a$, $\boldsymbol{\phi}_b$, $\boldsymbol{\phi}_c$, $\boldsymbol{\phi}_d$, $\boldsymbol{\phi}_e$, and $\boldsymbol{\phi}_f$, respectively. $\boldsymbol{\phi}_a$ and $\boldsymbol{\phi}_b$ are bias-independent; $\boldsymbol{\phi}_c$ and $\boldsymbol{\phi}_d$ are unconstrained bias-dependent; $\boldsymbol{\phi}_e$ and $\boldsymbol{\phi}_f$ are bias-dependent and constrained by the dc characteristics of the device. $\boldsymbol{\phi}_a$ and $\boldsymbol{\phi}_c$ appear in both the dc and ac models, whereas $\boldsymbol{\phi}_b$ and $\boldsymbol{\phi}_d$ affect only the ac small-signal equivalent circuit.

We use superscript k to indicate a different bias point and the corresponding model. Therefore, $\boldsymbol{\phi}_c^k$, $\boldsymbol{\phi}_d^k$, $\boldsymbol{\phi}_e^k$ and $\boldsymbol{\phi}_f^k$ belong to the model under the kth bias, whereas $\boldsymbol{\phi}_a$ and $\boldsymbol{\phi}_b$ remain unchanged for the different bias points.

We express the functional dependency of $\boldsymbol{\phi}_e$ and $\boldsymbol{\phi}_f$ on the bias by $\boldsymbol{\phi}_e^k = \boldsymbol{\phi}_e(\boldsymbol{\alpha}, \mathbf{v}^k)$ and $\boldsymbol{\phi}_f^k = \boldsymbol{\phi}_f(\boldsymbol{\beta}, \mathbf{v}^k)$, where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the coefficients of the constraints, and $\mathbf{v}^k = \mathbf{v}(\boldsymbol{\phi}_a, \boldsymbol{\phi}_c^k, \boldsymbol{\alpha})$ denotes the dc state variables (such as the voltages and currents of the dc circuit). The constraints on $\boldsymbol{\phi}_e^k$ and $\boldsymbol{\phi}_f^k$ may be derived from physical characteristics of the device. They may be introduced empirically to simulate the dc characteristics. These constraints reduce the degrees of freedom in modeling, since the number of variables in this group, namely $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, does not increase when more bias points are used. $\boldsymbol{\alpha}$ is part of the dc equivalent circuit but $\boldsymbol{\beta}$ is not.

Assume that the dc and ac measurements and the corresponding model responses are $\mathbf{S}_{dc}^k$ and $\mathbf{S}_{ac}^k(\omega_n)$, $\mathbf{F}_{dc}^k$ and $\mathbf{F}_{ac}^k(\omega_n)$, respectively, where

$$\begin{aligned} \mathbf{F}_{dc}^k &= \mathbf{F}_{dc}(\boldsymbol{\phi}_a, \boldsymbol{\phi}_c^k, \boldsymbol{\alpha}), \\ \mathbf{F}_{ac}^k(\omega_n) &= \mathbf{F}_{ac}(\boldsymbol{\phi}_a, \boldsymbol{\phi}_b, \boldsymbol{\phi}_c^k, \boldsymbol{\phi}_d^k, \boldsymbol{\phi}_e(\boldsymbol{\alpha}, \mathbf{v}^k), \boldsymbol{\phi}_f(\boldsymbol{\beta}, \mathbf{v}^k); \omega_n), \end{aligned} \tag{1}$$

and ω_n, n=1, 2, ..., N, is a set of frequency points. Thus, the error functions can be expressed as

$$\begin{aligned} e_{dcj}^k &= w_{dcj}^k(F_{dcj}^k - S_{dcj}^k), \quad j=1,2,\ldots,M_{dc}^k, \quad k=1,2,\ldots,K, \\ e_{acj}^k(\omega_n) &= w_{acj}^k(F_{acj}^k(\omega_n) - S_{acj}^k(\omega_n)), \\ & j=1,2,\ldots,M_{ac}^k, \quad n=1,2,\ldots,N, \quad k=1,2,\ldots,K, \end{aligned} \tag{2}$$

1988 IEEE MTT-S Digest

where M_{dc}^{k} and M_{ac}^{k} are the numbers of dc and ac measurements taken at the kth bias point, respectively. K is the total number of different bias points. w_{dcj}^{k} and w_{acj}^{k} are weighting factors.

To obtain a uniform set of error functions, we define

$$\begin{aligned} f_i &= e_{dcj}^{k}, \ \ j=1,2,...,M_{dc}^{k}, \ \ k=1,2,...,K, \ \ i \in J_{dc}, \\ f_i &= e_{acj}^{k}(\omega_n), \ j=1,...,M_{ac}^{k}, \ \ n=1,...,N, \ k=1,...,K, \ i \in J_{ac}, \end{aligned} \tag{3}$$

where J_{dc} and J_{ac} are sets containing unique indices for **f**. Then we can formulated the ℓ_1 modeling optimization problem

$$\underset{\boldsymbol{\phi}_a,\boldsymbol{\phi}_b,\boldsymbol{\phi}_c^k,\boldsymbol{\phi}_d^k,\boldsymbol{\alpha},\boldsymbol{\beta}}{\text{minimize}} \quad \{ \sum_{i\in J_{dc}} |f_i| \ + \sum_{i\in J_{ac}} |f_i| \}. \tag{4}$$

To calculate the model responses, we first solve the nonlinear dc circuit, and then, with known $\boldsymbol{\phi}_e(\boldsymbol{\alpha},\mathbf{v}^k)$ and $\boldsymbol{\phi}_f(\boldsymbol{\beta},\mathbf{v}^k)$, solve the linear ac circuit. The derivatives of the error functions can be obtained efficiently by adjoint analyses of both the dc and ac circuits.

SEQUENTIAL MODEL BUILDING

A device model, such as the FET model in Super-Compact [8], may have complicated topology and a comprehensive set of possible model parameters. In practice, we prefer a simplified model, provided the match between the model responses and the measurements is satisfactory. It not only simplifies the computation, but also increases the identifiability.

Approaches have been proposed (e.g., [9]) which optimize both the element values and the model topology. However, the topology optimization part of these approaches is entirely by trial and error and quite often has no physical justification.

For sequential model building, we apply the approach of Bandler and Zhang [3] to construct a decomposition dictionary to identify the interdependency between the model responses and parameters.

We start with a basic model structure, and sequentially add parameter(s) that according to the decomposition dictionary would most effectively improve the match between the model responses and the measurements. The decomposition dictionary may also reveal parameters that are impossible or very difficult to be identified from the available measurements. Such parameters could be kept fixed at standard values. They may even be eliminated from the model if they have little effect on the match between the model responses and the measurements.

Consider a function $f_j(x)$ and a parameter x_i. A measure of the degree of interdependency between x_i and f_j can be defined, following [3], as

$$C_{ij} = \sum_{r=1}^{L} | \frac{\partial f_j(x^r)}{\partial x_i} x_i^r |^p \tag{5}$$

where L is the number of points randomly chosen around x, x_i^r is a scaling factor, and p can be 1 or 2. (In the example discussed in the next section we choose p=1.) The decomposition dictionary is constructed by further grouping closely related functions

$$D_{it} = \sum_{j\in J_t} C_{ij} \tag{6}$$

where $J_1 \cup J_2 \cup .. \cup J_q = J_{dc} \cup J_{ac}$, and q is the number of function groups. For instance, we may designate all the error functions related to S_{11} to one function group. The relative magnitude of D_{it} indicates the relative degree of interdependency between parameter x_i and the t-th function group.

A FET EXAMPLE

The FET equivalent circuit model is shown in Fig. 1(a). The corresponding small-signal equivalent circuit is shown in Fig. 1(b). We use measurements made under three different biasing conditions (the same data has been considered by Bandler et al. [7]).

Following our earlier discussions and the assumptions in [7], we define bias-independent parameters as

$$\boldsymbol{\phi}_a = [R_g \ \ R_d]^T, \ \ \boldsymbol{\phi}_b = [L_g \ \ L_d \ \ L_s \ \ \tau]^T, \tag{7}$$

unconstrained bias-dependent parameters as

$$\boldsymbol{\phi}_c^k = [R_i^k \ \ R_s^k]^T, \ \ \boldsymbol{\phi}_d^k = [C_{dg}^k \ \ C_{ds}^k]^T \ \ k = 1, 2, 3, \tag{8}$$

and constrained bias-dependent parameters as

$$\boldsymbol{\phi}_e^k = [G_{ds}^k \ \ g_m^k]^T, \ \ \boldsymbol{\phi}_f^k = [C_{gs}^k], \ \ \ k = 1, 2, 3. \tag{9}$$

The dc constraints imposed on G_{ds}, g_m and C_{gs} are described by Materka and Kacprzak [5]:

$$\begin{aligned} i_f &= I_s[\exp(\alpha_s v_g) - 1], \\ i_r &= I_{sr}[\exp(\alpha_{sr} v_{dg}) - 1], \\ i_d &= I_{dss}(1 - v_g/V_p)^2 \tanh(\alpha_d v_d/(v_g - V_p)), \\ V_p &= V_{po} + \gamma \, v_d, \\ C_{gs} &= C_{go}(1 - v_g/V_{bi})^{-0.5}, \text{ for } v_g < 0.8V_{bi}, \end{aligned} \tag{10}$$

of which the optimizable coefficients are given by

$$\begin{aligned} \boldsymbol{\alpha} &= [I_s \ \ \alpha_s \ \ I_{sr} \ \ \alpha_{sr} \ \ I_{dss} \ \ \alpha_d \ \ V_{po} \ \ \gamma]^T, \\ \boldsymbol{\beta} &= [C_{go} \ \ V_{bi}]^T. \end{aligned} \tag{11}$$

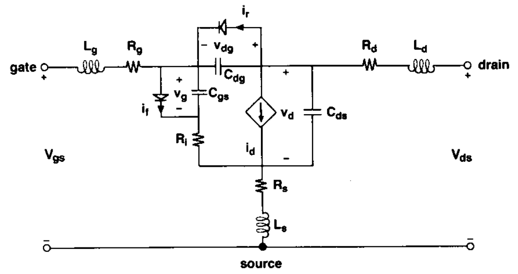
Altogether there are 28 parameters in $\boldsymbol{\phi}_a$, $\boldsymbol{\phi}_b$, $\boldsymbol{\phi}_c^k$ and $\boldsymbol{\phi}_d^k$ for k = 1, 2, 3, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

The error functions are defined according to (2), where K=3 for three different bias points; $M_{dc}^{k}=2$ corresponding to the dc measurements on the gate and source currents; $M_{ac}^{k}=8$ representing the real and imaginary parts of the S-parameters; and N=17 representing 17 frequency points from 2GHz to 18GHz, 1GHz apart. The weighting factors w_{dcj}^{k} and w_{acj}^{k} are properly chosen to balance the dc and ac error functions. The total number of error functions for this example is 414.

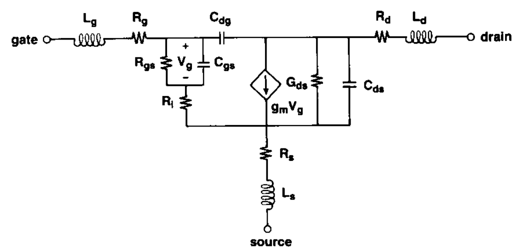
Case 1 At the starting point, we construct the decomposition dictionary which showed very small entries for I_s, α_s, I_{sr} and α_{sr}. An ℓ_1 optimization is performed, fixing $I_s=I_{sr}=0.5$nA, $\alpha_s=20$ and $\alpha_{sr}=1$. The resulting parameter values are listed in Table I. Table II shows the dc responses, and Fig. 2 depicts the ac responses at the solution for one bias point.

To check whether we should consider I_s, α_s, I_{sr} and α_{sr} as variables, we set up the dictionary at the solution, as shown in Table III. The fact that the entries for I_s, α_s, I_{sr} and α_{sr} remain very small confirms the validity of eliminating them as optimization variables. As a further verification, we attempted another optimization which included all possible variables. As expected, it did not improve the match between the model responses and the measurements.

Also from the dictionary we can see that the entries for R_g and R_d are very small, therefore their optimized values may not be reliable.



(a)



(b)

Fig. 1 (a) FET equivalent circuit model
(b) the corresponding small-signal equivalent circuit

TABLE I PARAMETER VALUES OF THE FET MODEL

Para.	Bias 1 start	Bias 1 solution	Bias 2 start	Bias 2 solution	Bias 3 start	Bias 3 solution
R_g^+	1.0	0.0119	1.0	0.0119	1.0	0.0119
R_d^+	1.0	0.0006	1.0	0.0006	1.0	0.0006
G_{ds}	*	0.0049	*	0.0058	*	0.0063
R_i	1.0	3.4731	1.0	4.2221	1.0	5.5954
R_s	1.0	0.5234	1.0	0.3675	1.0	0.2312
L_s	0.02	0.0107	0.02	0.0107	0.02	0.0107
C_{gs}	*	0.5929	*	0.3992	*	0.3333
C_{dg}	0.07	0.0287	0.07	0.0428	0.07	0.0533
C_{ds}	0.04	0.1958	0.04	0.1917	0.04	0.1905
g_m	*	0.0569	*	0.0437	*	0.0302
τ	7.0	3.6540	7.0	3.6540	7.0	3.6540
L_g	0.01	0.1257	0.01	0.1257	0.01	0.1257
L_d	0.01	0.0719	0.01	0.0719	0.01	0.0719

Coef.	start	solution	Coef.	start	solution
I_{dss}	0.2	0.1888	γ	-0.2	-0.3958
α_d	4.0	3.0523	C_o	1.0	0.6137
V_{po}	-4.0	-4.3453	V_{bi}	1.0	1.3011

See biasing conditions in Table II.
\+ values may not be reliable as the decomposition dictionary shows weak identifiability
* values determined by α, β and dc solution

To test the robustness of our approach, we randomly perturb the starting point by 20 to 200 percent and restart the optimization. All the variables converged to virtually the same solution.

TABLE II DC RESPONSES AND MEASUREMENTS

		Bias 1	Bias 2	Bias 3
		V_{gs} = 0V, V_{ds} = 4V	V_{gs} = −1.74V, V_{ds} = 4V	V_{gs} = −3.1V, V_{ds} = 4V
I_{gs}	assumed	0.0A	0.0A	0.0A
i_{gs}	response	-2.7×10^{-8}A	-1.5×10^{-7}A	-6.1×10^{-7}A
I_{ds}	measured	0.177A	0.092A	0.037A
i_{ds}	response	0.177A	0.092A	0.043A

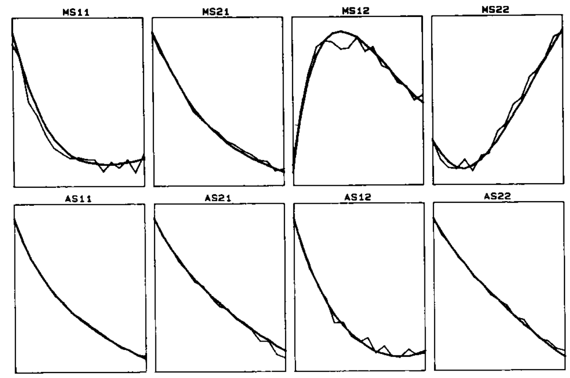


Fig. 2 The S-parameter match at the solution of Case 1 for V_{ds} = 4V and V_{gs} = 0V.

Case 2 To demonstrate the feasibility of sequential model building, we restart the modeling process with a simplified model which does not include L_g and L_d. Also, R_g, R_d, R_i, I_s, α_s, I_{sr} and α_{sr} are kept constant according to their relatively small entries in the decomposition dictionary. Fig. 3 depicts the model responses and the measurements at one bias point after the ℓ_1 optimization using this simplified model.

It is obvious from Fig. 3 that the worst match is for S_{11}. Our decomposition dictionary indicates that the most effective candidates for improving the match in S_{11} are R_g and L_g because of their larger entries under S_{11}. The result of a subsequent optimization which includes R_g and L_g as variables is shown in Fig. 4, from which a significant improvement on the match of S_{11} can be observed.

Further steps of sequential model building based on the decomposition dictionary include adding L_d to improve S_{22} and eventually converge to the same solution as in Case 1. All the intermediate results were exactly what we expected.

CONCLUSIONS

By introducing dc constraints and formulating the modeling process as a complete and integrated optimization problem, we have improved the uniqueness and reliability of the extracted model parameters. A sequential model building approach has been proposed based on a decomposition dictionary. It can be used to arrive at a suitable compromise between the simplicity and adequacy of the model.

A powerful ℓ_1 optimization technique has been employed in our algorithm, and all the required gradients have been provided through efficient adjoint analyses.

A FET modeling example has been described in detail to demonstrate the new approach.

TABLE III DECOMPOSITION DICTIONARY

Para.	I_{gs}	I_{ds}	S_{11}	S_{21}	S_{12}	S_{22}
R_g	0.00	0.00	0.02	0.01	0.01	0.00
L_g	-	-	15.	3.8	10.	1.1
L_d	-	-	0.32	1.6	4.6	9.2
L_s	-	-	0.91	0.24	16.	0.89
R_d	0.00	0.00	0.00	0.00	0.00	0.00
τ	-	-	1.0	6.3	1.5	2.6
C_{dg}^1	-	-	1.7	1.7	28.	4.4
C_{ds}^1	-	-	0.36	4.1	9.8	16.
R_i^1	0.00	0.00	1.4	0.54	3.5	0.16
R_s^1	0.00	0.55	0.64	0.39	6.6	0.53
C_{dg}^2	-	-	3.4	3.0	21.	6.6
C_{ds}^2	-	-	0.80	4.4	9.3	16.
R_i^2	0.00	0.00	1.6	0.60	2.1	0.24
R_s^2	0.00	0.15	0.20	0.18	1.8	0.13
C_{dg}^3	-	-	4.2	3.3	19.	6.9
C_{ds}^3	-	-	0.94	4.0	9.2	16.
R_i^3	0.00	0.00	2.1	0.61	2.1	0.26
R_s^3	0.00	0.03	0.07	0.08	0.74	0.09
I_s	0.00	0.00	0.00	0.00	0.00	0.00
α_s	0.00	0.00	0.00	0.00	0.00	0.00
I_{sr}	0.00	0.00	0.00	0.00	0.00	0.00
α_{sr}	0.00	0.00	0.00	0.00	0.00	0.00
I_{dss}	0.00	29.	4.2	33.	18.	42.
α_d	0.00	2.8	1.2	12.	11.	25.
V_{po}	0.00	13.	2.8	26.	12.	28.
γ	0.00	4.6	1.4	12.	7.1	19.
C_o	-	-	40.	26.	49.	9.9
V_{bi}	-	-	9.3	5.5	9.7	2.4

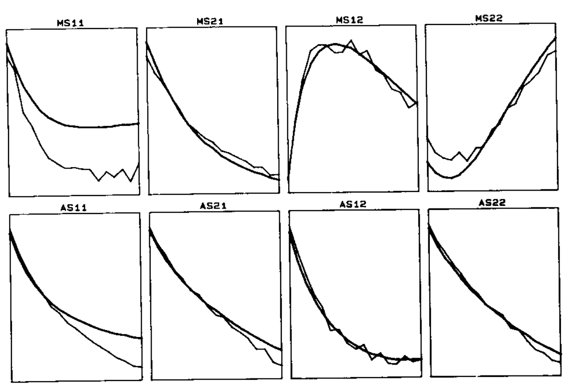


Fig. 3 The S-parameter match at the solution of Case 2 using a simplified model for V_{ds} = 4V and V_{gs} = 0V.

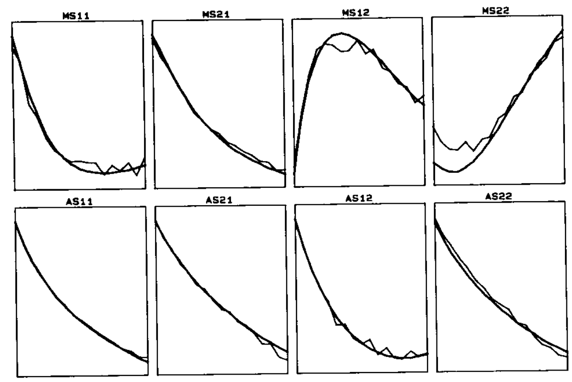


Fig. 4 The S-parameter match at the solution of Case 2 for V_{ds} = 4V and V_{gs} = 0V. R_g and L_g were included as optimization variables.

REFERENCES

[1] J.W. Bandler, S.H. Chen and S. Daijavad, "Microwave device modeling using efficient ℓ_1 optimization: a novel approach", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-34, 1986, pp. 1282-1293.

[2] J.W. Bandler and S.H. Chen, "Circuit optimization: the state of the art", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-36, 1988, pp. 424-443.

[3] J.W. Bandler and Q.J. Zhang, "An automatic decomposition approach to optimization of large microwave systems", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-35, 1987, pp. 1231-1239.

[4] W. Curtice, "GaAs MESFET modeling and nonlinear CAD", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-36, 1988, pp. 220-229.

[5] 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.

[6] J.W. Bandler, W. Kellermann and K. Madsen, "A nonlinear ℓ_1 optimization algorithm for design, modeling and diagnosis of networks", *IEEE Trans. Circuit and Systems*, vol. CAS-34, 1987, pp. 174-181.

[7] J.W. Bandler, S.H. Chen, S. Daijavad and K. Madsen, "Efficient optimization with integrated gradient approximations", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-36, 1988, pp. 444-455.

[8] Super-Compact User's Manual, Compact Software Inc., Patterson, NJ 07504, April 1987.

[9] M. Dowson, "Computer-aided design of equivalent circuit models for microwave frequencies", *Computer-aided Design*, vol. 17, no. 8, 1985, pp. 353-362.