INTEGRATED MODEL PARAMETER EXTRACTION USING LARGE-SCALE OPTIMIZATION CONCEPTS

OSA-88-MT-5-R

April 4, 1988

INTEGRATED MODEL PARAMETER EXTRACTION USING LARGE-SCALE OPTIMIZATION CONCEPTS

J.W. Bandler, Fellow, IEEE, S.H. Chen, Student Member, IEEE,
S. Ye, and Q.J. Zhang, Student Member, IEEE

Abstract This paper presents a robust approach to model parameter extraction. The approach not only attempts to match dc and ac measurements under different bias conditions simultaneously, but also employs the dc characteristics of the device as constraints on bias-dependent parameters, hence improving the uniqueness and reliability of the solution. The approach is an expansion of the hierarchical modeling techniques recently proposed by Bandler and Chen. Based on Bandler and Zhang's automatic decomposition concepts for large-scale optimization, a sequential model building method is proposed which, combined with powerful ℓ_1 optimization techniques, can be used to establish a model with simple topology and sufficient accuracy.

Practical FET models proposed by Materka and Kacprzak and by Curtice and Ettenberg are used to illustrate our formulation. A detailed numerical example based on the Materka and Kacprzak model is presented which has up to 28 optimization variables and 414 nonlinear error functions. The results show that a unique solution can be reached even after perturbing the original starting point (initial model parameter values) by 20 to 200 percent. The results have also shown the effectiveness of applying the sequential model building method to the FET modeling problem.

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 Grant A7239.

J.W. Bandler, S.H. Chen and Q.J. Zhang are with Optimization Systems Associates Inc., 163 Watson's Lane, Dundas, Ontario, Canada L9H 6L1. J.W. Bandler is also with and S. Ye is with the Simulation Optimization Systems Research Laboratory and the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada L8S 4L7.

I. INTRODUCTION

Model parameter extraction, i.e., the determination of equivalent circuit parameters from dc, rf, and microwave measurements on devices (such as FETs), is of fundamental importance to microwave circuit designers. 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., [1-3]) which separate the dc, low frequency and high frequency measurements and divide the model parameters into corresponding subsets. This defines a set of subproblems to be solved sequentially. Such a sequentially decoupled solution, however, may not be reliable: a parameter determined solely from dc measurements may not be suitable for the purpose of microwave simulation, and the information contained in ac measurements is not fully utilized.

The multi-circuit algorithm [4,5] can improve the uniqueness of the solution by simultaneously processing multiple sets of S-parameter measurements made under different bias conditions. However, the authors [4,5] assumed for computational purposes that the model parameters were either completely bias-independent or arbitrarily bias-dependent.

The approach presented in this paper 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. This enables us to use more efficiently the information contained in dc and non-zero frequency measurements and reduce the degrees of freedom by imposing constraints on bias-dependent parameters. In this way we aim at improving the uniqueness and reliability of the solution.

Bandler and Zhang [6] have proposed a decomposition dictionary to reveal the interdependency between functions and their variables. In this paper, such a dictionary and the powerful ℓ_1 optimization algorithm [7] are integrated to explore the relations between the model responses and model parameters during the modeling process, so that possible model defects could be overcome sequentially. In other words, we start the modeling process with the simplest model structure, subsequently adding elements according to the ℓ_1 optimization result and the dictionary for a better match between the model responses and the measurements.

In Section II, through a simple circuit example we demonstrate the feasibility and usefulness of integrating dc and ac modeling in one optimization problem. In Section III, general and abstract definitions for the model parameters are given. The definitions are illustrated by examples of significant interest, namely the Materka and Kacprzak FET model [2] and the Curtice and Ettenberg FET model [8]. The modeling optimization problem with both dc and ac responses is formulated in Section IV. In Section V, we present the sequential model building approach. In Section VI, a FET modeling example using the Materka and Kacprzak model is described in detail to demonstrate our new approach.

II. A SIMPLE CIRCUIT EXAMPLE

As a simple example to illustrate that combining dc and ac modeling is both feasible and useful, let us examine the linear RC circuit shown in Fig. 1. The unknown parameters are $\boldsymbol{\phi} = [R_1\ R_2\ C]^T$. R_3 is assumed to be a known resistor. We also assume the responses to be the dc current I, under dc excitation $V_1 = V_{dc}$, as

$$I = \frac{V_{dc}}{R_1 + R_2} \tag{1}$$

and ac (complex) voltage V_2, under ac excitation $V_1 = V_{ac}$, as

$$V_2 = \frac{V_{ac}R_2R_3\ s\ C}{sC(R_1R_2+R_1R_3+R_2R_3)+R_1+R_2}\ , \tag{2}$$

where s denotes the complex frequency variable.

It is obvious that we cannot distinguish R_1 and R_2 if only the dc response I is used. It can also be verified that if only the ac response V_2 is taken, we cannot uniquely determine $\boldsymbol{\phi}$ either, no matter how many frequency points are applied.

It can be proved, however, that $\boldsymbol{\phi}$ will be uniquely determined when we utilize both dc and ac measurements simultaneously, i.e., to match the dc response and ac response to the corresponding measurements at the same time. (The detailed proof of this observation is provided in Appendix A.)

III. CLASSIFICATION OF MODEL PARAMETERS

A. The general case

In general, consider a device model with its equivalent circuit. The model parameters can be 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, where $\boldsymbol{\phi}_a$ and $\boldsymbol{\phi}_b$ are bias-independent, $\boldsymbol{\phi}_c$ and $\boldsymbol{\phi}_d$ are unconstrained bias-dependent, and $\boldsymbol{\phi}_e$ and $\boldsymbol{\phi}_f$ are constrained bias-dependent. $\boldsymbol{\phi}_a$ 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 device 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 different bias points.

We express the functional dependency of $\boldsymbol{\phi}_e$ and $\boldsymbol{\phi}_f$ on the bias condition by $\boldsymbol{\phi}_e^k = \boldsymbol{\phi}_e(\boldsymbol{\alpha},\mathbf{v}^k)$ and $\boldsymbol{\phi}_f^k = \boldsymbol{\phi}_f(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{v}^k)$, 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). $\boldsymbol{\alpha}$ affects the dc equivalent circuit but $\boldsymbol{\beta}$ does not.

Table I summarizes the foregoing definitions.

This categorization stems from the consideration of the physical device and a feasible model. It is clear that we need $\boldsymbol{\phi}_a$ and $\boldsymbol{\phi}_b$ to represent the parameters which do not or almost not vary with the bias conditions, such as package capacitance and lead inductance of an FET. We need $\boldsymbol{\phi}_c^k$ and $\boldsymbol{\phi}_d^k$ to represent those bias-dependent parameters whose functional bias dependency expressions may not be known or available; on the other hand $\boldsymbol{\phi}_c^k$ and $\boldsymbol{\phi}_d^k$ may be used to test or investigate the functional bias-dependent properties of the model parameters.

Introducing $\boldsymbol{\phi}_e(\boldsymbol{\alpha},\mathbf{v}^k)$ and $\boldsymbol{\phi}_f(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{v}^k)$ allows us to describe other bias-dependent parameters whose bias-dependent properties can be expressed by functions or, as we refer to them, constraints. Such constraints may be derived from physical characteristics of the device. They may be introduced empirically to simulate the pattern of the dc characteristic curves. They may also include mathematical expressions, such as polynomials. 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, so that the uniqueness of the solution can be improved.

Our classification of the model parameters is consistent with the hierarchical parameter descriptions of Bandler and Chen [5]. From the definitions presented above, for example, we can see that for the ac responses, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are low level parameters compared with $\boldsymbol{\phi}_a$, $\boldsymbol{\phi}_b$, $\boldsymbol{\phi}_c$, $\boldsymbol{\phi}_d$, $\boldsymbol{\phi}_e$ and $\boldsymbol{\phi}_f$. However, we should notice that a parameter can appear as a low level and high level parameter simultaneously. For example, $\boldsymbol{\phi}_e$ is at same level as $\boldsymbol{\phi}_a$ but it depends on $\boldsymbol{\phi}_a$ as well.

B. Two practical FET device examples

To illustrate the definitions presented in the previous subsection, we consider a typical nonlinear FET model proposed by Materka and Kacprzak [2]. The model and its corresponding small-signal equivalent circuit are shown in Figs. 2(a) and 2(b), respectively.

In the Materka and Kacprzak model, there are three bias-dependent current sources i_f, i_r and i_d, (see Fig. 2(a)), which are defined as [2]

$$
\begin{aligned}
i_f &= I_s\,[\exp(\alpha_s v_g) - 1],\\
i_r &= I_{sr}\,[\exp(\alpha_{sr} v_{dg}) - 1],\\
i_d &= I_{dss}\,(1 - \frac{v_g}{V_p})^2 \tanh(\frac{\alpha_d v_d}{v_g - V_p}),
\end{aligned}
\qquad (3)
$$

where

$$V_p = V_{po} + \gamma v_d,$$

and where I_s, α_s, I_{sr}, α_{sr}, I_{dss}, α_d, V_{po} and γ are parameters to be determined. Three other bias-dependent parameters G_{ds}, g_m and C_{gs} (see also Fig. 2(b)) are constrained by [2]

$$G_{ds} = \frac{\partial i_d}{\partial v_d},$$

$$g_m = \frac{\partial i_d}{\partial v_g}, \tag{4}$$

$$C_{gs} = C_{go}(1 - \frac{v_g}{V_{bi}})^{-0.5}, \text{ for } v_g < 0.8V_{bi},$$

where C_{go} and V_{bi} are also parameters to be determined.

Table II gives clear classifications for all the parameters of the Materka and Kacprzak model.

We have also considered another typical nonlinear FET model proposed by Curtice and Ettenberg [8], as shown in Fig. 3(a). Its small-signal equivalent circuit is shown in Fig. 3(b). Following the considerations and the notation of [8], I_{gs} is a function of $\{R_F, V_{bi}\}$, I_{dg} is a function of $\{R_1, R_2, V_{B0}\}$, I_{ds}, g_{ds} and g_m are functions of $\{A_0, A_1, A_2, A_3, \gamma, \beta, V^0_{out}\}$, C_{gs} is a function of $\{V_{BI}\}$, and τ is a function of $\{A_5\}$. The classifications of the parameters are listed in the last column of Table II. (For details of the Curtice and Ettenberg model, see [8].)

IV. MULTI-BIAS DC AND AC MODELING OPTIMIZATION

Assume that the dc and ac measurements are S^k_{dc} and $S^k_{ac}(\omega_n)$, respectively, where ω_n, n = 1, 2, ..., N, is a set of frequency points. Correspondingly, we assume

$$F^k_{dc} = F_{dc}(\boldsymbol{\phi}_a, \boldsymbol{\phi}^k_c, \boldsymbol{\alpha}) \tag{5}$$

as the dc model response, and

$$F^k_{ac}(\omega_n) = F_{ac}(\boldsymbol{\phi}_a, \boldsymbol{\phi}_b, \boldsymbol{\phi}^k_c, \boldsymbol{\phi}^k_d, \boldsymbol{\phi}_e(\boldsymbol{\alpha}, v^k), \boldsymbol{\phi}_f(\boldsymbol{\alpha}, \boldsymbol{\beta}, v^k); \omega_n) \tag{6}$$

as the ac model response. Thus, the error functions corresponding to the dc model responses can be expressed as

$$e^k_{dcj} = w^k_{dcj}\,(F^k_{dcj} - S^k_{dcj}),\quad j = 1, 2, ..., M^k_{dc},\ k \in K_{dc} \tag{7}$$

where w^k_{dcj} is the weighting factor, M^k_{dc} is the number of dc measurements taken at the kth bias point, and K_{dc} is the set of bias point at which dc measurements are taken. The error functions corresponding to the ac model responses can be expressed as

$$e^k_{acj}(\omega_n) = w^k_{acj}\,[F^k_{acj}(\omega_n) - S^k_{acj}(\omega_n)], \tag{8}$$
$$j = 1, 2, ..., M^k_{ac},\ n = 1, 2, ..., N,\ k \in K_{ac},$$

where w^k_{acj} is the weighting factor, M^k_{ac} is the number of ac measurements taken at the kth bias point, and K_{ac} is the set of bias point at which ac measurements are taken.

If we use K to indicate the set of all bias points, then

$$K = K_{dc} \cup K_{ac} = \{1, 2, ..., K_{bias}\}. \tag{9}$$

Usually M^k_{dc} could be the same for different k, $k \in K_{dc}$, such as the number of dc current responses at different bias conditions. Similarly M^k_{ac} could be the same for different k, $k \in K_{ac}$, such as the number of S-parameter responses.

To obtain a uniform set of error functions, we define

$$f_i = e^k_{dcj},\quad j = 1, 2, ..., M^k_{dc},\ k \in K_{dc},\ i \in J_{dc}, \tag{10}$$

and

$$f_i = e^k_{acj}(\omega_n),\quad j = 1, 2, ..., M^k_{ac},\ n = 1, 2, ..., N,\ k \in K_{ac},\ i \in J_{ac}, \tag{11}$$

where $J_{dc} = \{1, 2, ..., M_1\}$, M_1 is the total number of dc measurements, $J_{ac} = \{M_1+1, M_1+2, ..., M_2\}$, and M_2 is the total number of measurements. Then we can formulate the ℓ_1 modeling optimization problem

$$\text{minimize}\quad \{ \sum_{i \in J_{dc}} |f_i| + \sum_{i \in J_{ac}} |f_i| \}, \tag{12}$$

where the optimization variables are $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\phi}_a$, $\boldsymbol{\phi}_b$, $\boldsymbol{\phi}^k_c$ for $k \in K$, and $\boldsymbol{\phi}^k_d$ for $k \in K_{ac}$, since $\boldsymbol{\phi}^k_c$ is required for calculating both dc and ac responses, whereas $\boldsymbol{\phi}^k_d$ is only required for calculating ac responses.

In order to calculate the model responses, we first solve the nonlinear dc circuit of the model for $\boldsymbol{\phi}_a$, $\boldsymbol{\phi}_c^k$ and $\boldsymbol{\alpha}$, $k \in K$, so that $\mathbf{F}_{dc}^k$, if $k \in K_{dc}$, can be determined. If $k \in K_{ac}$, $\boldsymbol{\phi}_e(\boldsymbol{\alpha},\mathbf{v}^k)$ and $\boldsymbol{\phi}_f(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{v}^k)$ are calculated with $\mathbf{v}^k$ obtained from the dc solution. And then $\mathbf{F}_{ac}^k(\omega_n)$, $n = 1, 2, ..., N$, can be determined.

The derivatives of the error functions required by the optimization can be obtained by the perturbation method. However, since the equivalent circuit of the device model is usually not very complicated, it is both feasible and efficient to get them analytically by adjoint analyses. The details of the analytic derivative calculations are discussed in Appendix B.

V. SEQUENTIAL MODEL BUILDING

A device model, such as the FET model in Super-Compact [9], may have a complicated topology and a comprehensive set of possible model parameters. In practice, we prefer a simplified model, provided that 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., [10-11]) 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 start with a simple basic model structure, and sequentially add parameter(s) that would most effectively improve the match between the model responses and the measurements, where we assume that a comprehensive model which is physically meaningful is available. The iterative process continues until the match is satisfactory or no more parameter could be added. In order to find out the relationship between the model responses and parameters, we have applied the decomposition approach of Bandler and Zhang [6] to construct a decomposition dictionary to identify the interdependency between the model responses and parameters.

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 [6], as

$$C_{ij} = \sum_{r=1}^{L} \left| \frac{\partial f_j(x^r)}{\partial x_i} x_i^0 \right|^p \tag{13}$$

where L is the number of points randomly chosen around x, x_i^0 is a scaling factor, and p can be 1 or 2. (In the example discussed in the next section we will choose p = 1.) The decomposition dictionary is constructed by further grouping closely related functions

$$D_{it} = \sum_{j \in J_t} C_{ij} \tag{14}$$

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 the complex S parameter 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.

By virtue of the ℓ_1 optimization algorithm [7], which has the very desireable feature of isolating large errors among all the error functions, the sequential model building procedure can practically be implemented: during the modeling process the ℓ_1 solution and the corresponding decomposition dictionary at a specific model structure can indicate the most appropriate element(s) to be included in the model if the match has not been satisfactory. (See Case 2 of the example in the next section.)

The decomposition dictionary may reveal parameters that are impossible or very difficult to be identified from the available measurements, i.e, if the dictionary entries corresponding to a parameter are very small, this parameter may be very insensitive to any functions. 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.

VI. A FET EXAMPLE

Consider again the Materka and Kacprzak model discussed in Section III-B. The FET equivalent circuit model is shown in Fig. 2(a) and the corresponding small-signal equivalent circuit is shown in Fig. 2(b). We use measurements made under three different bias conditions (the same data has been considered by Bandler et al. [12]).

Following the assumptions in [12], we will use the classification of the parameters under **Model 1*** in Table II, however, we ignore the package parasitics L_{pg}, L_{pd}, C_{pg} and C_{pd}. Since there are three bias points, altogether we have 28 optimization variables 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 units of the related parameters are listed in Table III.

The error functions are defined according to (7) and (8), where $K_{dc} = K_{ac} = \{1, 2, 3\}$ 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 nonlinear error functions for this example is 414.

At each bias point, we use Powell's algorithm [13] to solve the nonlinear dc equivalent circuit. The adjoint network analysis technique is applied to efficiently calculate the sensitivities of both dc and ac equivalent circuits.

Three cases are discussed as follows. In Case 1, we will show the robustness of the modeling approach proposed in Section IV. In Case 2, an experiment will be shown to demonstrate the feasibility of the sequential model building procedure in Section V. A similar experiment will be discussed briefly in Case 3 with a different scaling factor in (13).

Case 1: At the starting point, we construct the decomposition dictionary. The scaling factor x_i^0 in (13) for this dictionary is chosen to be x_i^r which corresponds to the exponential transformation on the variables used by the optimization. This dictionary shows 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, $\alpha_s = 20\text{V}^{-1}$ and $\alpha_{sr} = 1\text{V}^{-1}$. The resulting parameter values are listed in Table IV. Table V shows the dc responses, and Fig. 4 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 VI. 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.

The insensitivity of I_s, α_s, I_{sr} and α_{sr} is, in fact, expected, since it is known that special bias conditions are needed in order to effectively determine the forward-biasing and break-down properties of the FET [2].

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

Case 2: Similar to Case 1, the decomposition dictionaries in this case are constructed by choosing the scaling factor x_i^0 in (13) to be x_i^r which corresponds to the exponential transformation on the variables used by the optimization. To demonstrate the feasibility of the sequential model building procedure, 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. 5 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. 5 that the worst match is for S_{11}. According to the decomposition dictionary at this stage, as shown in Table VII, 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. 6, 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 R_d and L_d to improve S_{22} and eventually converge to the same solution as in Case 1.

By such a sequential model building, we have obtained a clear view of the relationship between a model parameter and the model responses, and we have the ability to avoid possible redundant model parameters. If the match between the model responses and measurements are sufficiently good, we do not have to include more optimizable parameters even if there are still some left.

Case 3: We also conducted an experiment where the decomposition dictionary was constructed by setting the scaling factor x_i^0 in (13) to 1, which corresponds to the sensitivities of the error functions w.r.t. the actual parameters.

At the starting point, we construct the dictionary with respect to all the possible variables. The actual variables used first in the optimization are those whose entries in the dictionary are relatively large, and other variables are kept constant. Then each time when an optimization is completed but the result is not satisfactory, we check the match between the model responses and measurements, and select new variable(s) according to the updated dictionary which would most effectively improve the match.

Following such a procedure, results similar to Case 2 were clearly observed. However, also observed from this experiment is that the parameters first chosen as optimization variables, i.e., the parameters whose entries are dominant in the decomposition dictionary, appeared to stay quite close to the first solution in the subsequent optimizations. Therefore, alternative decompositions for the optimization problem could be investigated.

VII. CONCLUSIONS

By introducing dc constraints and formulating the modeling process as a complete and integrated optimization problem, i.e., including simultaneously the dc and ac responses, 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, which is essential to the implementation of the sequential model building, has been employed in our algorithm. All the required gradients have been provided through efficient adjoint analyses.

Practical FET models have been considered. A FET modeling example using the Materka and Kacprzak model has been described in detail which clearly demonstrates the advantages of the new approach.

It should be noted, however, that when dc characteristics are used as constraints, they should be compatible with the actual device to be modeled, otherwise inappropriate dc constraints could cause large intrinsic discrepancies between the model responses and measurements.

As to the prospects of the approach proposed in this paper, we can see that

(1) The model parameters extracted can be used directly by the harmonic balance analysis.

(2) We can establish a more reliable small-signal model with dc constraints considered.

(3) The approach is applicable to other device modeling problems since it is quite general.

(4) The sequential model building procedure is particulary promising.

ACKNOWLEDGEMENT

The first author would like to acknowledge Dr. R.A. Pucel of Raytheon Company, Research Division, Lexington, MA, for many helpful discussions on device modeling for CAD.

REFERENCES

[1] H. Kondoh, "An accurate FET modelling from measured S-parameters", *IEEE Int. Microwave Symp. Digest* (Baltimore, MD), 1986, pp. 377-380.

[2] 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.

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

[4] 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.

[5] 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.

[6] 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.

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

[8] W.R. Curtice and M. Ettenberg, "A nonlinear GaAs FET model for use in the design of output circuits for power amplifiers", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-33, 1985, pp. 1383-1394.

[9] *SUPER-COMPACT User's Manual*, Compact Software Inc., Patterson, NJ 07504, April 1987.

[10] O.P.D. Cutteridge and Y.S. Zhang, "Computer-aided linear modelling of high-frequency bipolar transistors", *Proc. IEE*, vol. 131, part G, 1984, pp. 252-260.

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

[12] 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.

[13] P. Rabinowitz, *Numerical Methods for Nonlinear Algebraic Equations*. New York: Gordon and Breach Science Publishers, 1970, pp. 87-161.

[14] J.W. Bandler and A.E. Salama, "Fault diagnosis of analog circuits", *Proc. IEEE*, vol. 73, 1985, pp. 1279-1325.

[15] L.O. Chua and P. Lin, *Computer-aided Analysis of Electronic Circuits: Algorithms and Computational Techniques*. Englewood Cliffs, N.J.: Prentice-Hall, 1975.

APPENDIX A

VERIFICATION OF THE IDENTIFIABILITY OF THE RC CIRCUIT

The RC circuit under consideration is shown in Fig. 1. Similar to the derivations in [4], we use the concept of analog circuit fault diagnosis [14] to verify the identifiability of R_1, R_2 and C in the circuit. Briefly, given a complex-valued vector of responses $\mathbf{h}(\boldsymbol{\phi}) = [h_1(\boldsymbol{\phi}) \ldots h_m(\boldsymbol{\phi})]^T$, where $\boldsymbol{\phi} = [\phi_1 \ldots \phi_n]^T$, the measure of identifiability of $\boldsymbol{\phi}$ is determined by testing the rank of the m×n Jacobian matrix

$$\mathbf{J} = [\nabla \mathbf{h}^T(\boldsymbol{\phi})]^T, \tag{A1}$$

where ∇ is the partial derivative operator $\partial/\partial\boldsymbol{\phi}$. If the rank of matrix $\mathbf{J}$ is less than n, then $\boldsymbol{\phi}$ will not be uniquely identifiable from $\mathbf{h}$.

A. Only a dc response: The dc response I is

$$I = \frac{V_{dc}}{R_1 + R_2}. \tag{A2}$$

The corresponding Jacobian matrix is

$$\mathbf{J}_{dc} = [\ -V_{dc}/(R_1+R_2)^2 \quad -V_{dc}/(R_1+R_2)^2\ ]\ . \tag{A3}$$

It is clear that rank $\mathbf{J}_{dc} = 1$. Therefore, R_1 and R_2 are not identifiable from the dc response I. This result is also straightforward intuitively.

B. Only an ac response: The ac response V_2 is calculate as

$$V_2 = \frac{V_{ac} R_2 R_3\ s\ C}{sC(R_1R_2+R_1R_3+R_2R_3)+R_1+R_2}\ . \tag{A4}$$

The corresponding Jacobian matrix

$$\mathbf{J}_{ac} = \begin{bmatrix} -H[s_1C(R_2+R_3)+1] & H[R_1(s_1CR_3+1)/R_2] & H[(R_1+R_2)/C] \\ \vdots & \vdots & \vdots \\ -H[s_mC(R_2+R_3)+1] & H[R_1(s_mCR_3+1)/R_2] & H[(R_1+R_2)/C] \end{bmatrix} \tag{A5}$$

where s_i, i = 1, ..., m, indicate different frequencies, and

$$H = \frac{V_{ac} R_2 R_3\ s\ C}{[sC(R_1R_2+R_1R_3+R_2R_3)+R_1+R_2]^2}\ .$$

Denoting the three columns of $\mathbf{J}_{ac}$ by $\mathbf{J}_1$, $\mathbf{J}_2$ and $\mathbf{J}_3$, we can obtain

$$\frac{(R_2+R_3)}{R_1}\mathbf{J}_2 + \frac{R_3}{R_2}\mathbf{J}_1 - \frac{C}{(R_1+R_2)}\mathbf{J}_3 = \mathbf{0} \tag{A6}$$

which means that the rank of $\mathbf{J}_{ac}$ is less than 3, no matter how many frequency points are used. Hence we can not uniquely determine ϕ from the response V_2.

C. Combined dc and ac responses: When we consider both dc and ac responses simultaneously then, combining (A3) and (A5), the Jacobian matrix becomes

$$\mathbf{J} = \begin{bmatrix} -V_{dc}/(R_1+R_2)^2 & -V_{dc}/(R_1+R_2)^2 & 0 \\ -H[s_1C(R_2+R_3)+1] & H[R_1(s_1CR_3+1)/R_2] & H[(R_1+R_2)/C] \\ \vdots & \vdots & \vdots \\ -H[s_mC(R_2+R_3)+1] & H[R_1(s_mCR_3+1)/R_2] & H[(R_1+R_2)/C] \end{bmatrix} \tag{A7}$$

which is of full column rank. This indicates that ϕ is identifiable from the response h $= [I(\phi)\ V_2(\phi,s_1)\ \ldots\ V_2(\phi,s_m)]^T$.

All the three situations discussed above have been numerically proved.

APPENDIX B

DERIVATIVE COMPUTATIONS OF THE MODEL RESPONSES

For $j \in J_{dc}$, f_j is a function of $\boldsymbol{\phi}_a$, $\boldsymbol{\phi}_c^k$ and $\boldsymbol{\alpha}$, $k \in K_{dc}$. The corresponding $\partial f_j/\partial \boldsymbol{\phi}_a$, $\partial f_j/\partial \boldsymbol{\phi}_c^k$, and $\partial f_j/\partial \boldsymbol{\alpha}$ can be derived by nonlinear dc adjoint analysis [15].

For $j \in J_{ac}$, we know that

$$f_j = f_j(\boldsymbol{\phi}_a, \boldsymbol{\phi}_b, \boldsymbol{\phi}_c^k, \boldsymbol{\phi}_d^k, \boldsymbol{\phi}_e(\boldsymbol{\alpha}, \mathbf{v}^k), \boldsymbol{\phi}_f(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{v}^k)),\ k \in K_{ac} \tag{B1}$$

where the true variables are $\boldsymbol{\phi}_a$, $\boldsymbol{\phi}_b$, $\boldsymbol{\phi}_c^k$, $\boldsymbol{\phi}_d^k$, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, and $\mathbf{v}^k$ was defined in Section III-A. Therefore, we can use the chain rule to obtain the required derivatives

$$\begin{aligned} \frac{\partial f_j}{\partial \phi_{ai}} &= \frac{\partial f_j}{\partial \phi_{ai}} + \frac{\partial \mathbf{v}^{kT}}{\partial \phi_{ai}} \frac{\partial \boldsymbol{\phi}_e^{kT}}{\partial \mathbf{v}^k} \frac{\partial f_j}{\partial \boldsymbol{\phi}_e^k} + \frac{\partial \mathbf{v}^{kT}}{\partial \phi_{ai}} \frac{\partial \boldsymbol{\phi}_f^{kT}}{\partial \mathbf{v}^k} \frac{\partial f_j}{\partial \boldsymbol{\phi}_f^k} \\ \frac{\partial f_j}{\partial \phi_{bi}} &= \frac{\partial f_j}{\partial \phi_{bi}} \end{aligned} \tag{B2}$$

$$\begin{aligned} \frac{\partial f_j}{\partial \phi_{ci}^k} &= \frac{\partial f_j}{\partial \phi_{ci}^k} + \frac{\partial \mathbf{v}^{kT}}{\partial \phi_{ci}^k} \frac{\partial \boldsymbol{\phi}_e^{kT}}{\partial \mathbf{v}^k} \frac{\partial f_j}{\partial \boldsymbol{\phi}_e^k} + \frac{\partial \mathbf{v}^{kT}}{\partial \phi_{ci}^k} \frac{\partial \boldsymbol{\phi}_f^{kT}}{\partial \mathbf{v}^k} \frac{\partial f_j}{\partial \boldsymbol{\phi}_f^k} \\ \frac{\partial f_j}{\partial \phi_{di}^k} &= \frac{\partial f_j}{\partial \phi_{di}^k} \end{aligned} \tag{B3}$$

$$\begin{aligned} \frac{\partial f_j}{\partial \alpha_i} &= \frac{\partial \boldsymbol{\phi}_e^{kT}}{\partial \alpha_i} \frac{\partial f_j}{\partial \boldsymbol{\phi}_e^k} + \frac{\partial \mathbf{v}^{kT}}{\partial \alpha_i} \frac{\partial \boldsymbol{\phi}_e^{kT}}{\partial \mathbf{v}^k} \frac{\partial f_j}{\partial \boldsymbol{\phi}_e^k} + \frac{\partial \boldsymbol{\phi}_f^{kT}}{\partial \alpha_i} \frac{\partial f_j}{\partial \boldsymbol{\phi}_f^k} + \frac{\partial \mathbf{v}^{kT}}{\partial \alpha_i} \frac{\partial \boldsymbol{\phi}_f^{kT}}{\partial \mathbf{v}^k} \frac{\partial f_j}{\partial \boldsymbol{\phi}_f^k} \\ \frac{\partial f_j}{\partial \beta_i} &= \frac{\partial \boldsymbol{\phi}_f^{kT}}{\partial \beta_i} \frac{\partial f_j}{\partial \boldsymbol{\phi}_f^k} \end{aligned} \tag{B4}$$

where the superscript T stands for transposition, the derivative of f_j with respect to $\boldsymbol{\phi}_a$, $\boldsymbol{\phi}_b$, $\boldsymbol{\phi}_e$, $\boldsymbol{\phi}_f$, $\boldsymbol{\phi}_c^k$ and $\boldsymbol{\phi}_d^k$ for $k \in K_{ac}$ on the right hand side of (B2)-(B4) can be obtained by standard ac adjoint analysis, while the derivative of $\mathbf{v}^k$ with respect to $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\phi}_a$ and $\boldsymbol{\phi}_c^k$ for $k \in K_{ac}$ can be obtained by nonlinear dc adjoint analysis [15].

TABLE I

DEFINITIONS OF THE MODEL PARAMETERS

Category	Notation	Brief definition
Bias-independent	$\boldsymbol{\phi}_a$	affect dc and ac circuits
	$\boldsymbol{\phi}_b$	affect ac circuit
Unconstrained bias-dependent	$\boldsymbol{\phi}_c^k$	affect dc and ac circuits on the kth bias condition
	$\boldsymbol{\phi}_d^k$	affect ac circuit on the kth bias condition
Constrained bias-dependent	$\boldsymbol{\phi}_e(\boldsymbol{\alpha},\mathbf{v}^k)$	$\boldsymbol{\alpha}$ affects dc and ac circuits
	$\boldsymbol{\phi}_f(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{v}^k)$	$\boldsymbol{\beta}$ affects ac circuit only

$\mathbf{v}^k = \mathbf{v}(\boldsymbol{\phi}_a,\boldsymbol{\phi}_c^k,\boldsymbol{\alpha})$ denotes the dc state variables (such as the voltages and currents) under the kth bias condition.

TABLE II

PARAMETER DEFINITIONS FOR FET MODELS

Category	Subset	Parameters		
		Model 1	**Model 1***	**Model 2**
bias-independent	ϕ_a	R_g, R_d, R_s, R_i	R_g, R_d	R_g, R_d, R_s
	ϕ_b	L_g, L_d, L_s, L_{pg} L_{pd}, C_{pg}, C_{pd} C_{dg}, C_{ds}, τ	L_g, L_d, L_s, L_{pg} L_{pd}, C_{pg}, C_{pd} τ	R_{in}, C_{ds}
unconstrained bias-dependent	ϕ_c^k		R_i^k, R_s^k	R_{ds}^k
	ϕ_d^k		C_{dg}^k, C_{ds}^k	C_{dg}^k
constrained bias-dependent	ϕ_e	G_{ds}, g_m	G_{ds}, g_m	g_{ds}, g_m
	ϕ_f	C_{gs}	C_{gs}	C_{gs}, τ
	α	I_s, α_s, I_{sr}, α_{sr} I_{dss}, α_d, V_{po}, γ	I_s, α_s, I_{sr}, α_{sr} I_{dss}, α_d, V_{po}, γ	A_0, A_1, A_2, A_3 γ, β, V_{out}^0 R_1, R_2, V_{B0} V_{bi}, R_F
	β	C_{go}, V_{bi}	C_{go}, V_{bi}	V_{BI}, A_5

Notes:

(1) The parameters under **Model 1** are defined according to Materka and Kacprzak [2].

(2) The parameters under **Model 1*** are the same as those in **Model 1** except that we assume R_i, R_s, C_{dg} and C_{ds} to be bias-dependent but we do not enforce their bias-dependent characteristics.

(3) The parameters under **Model 2** are defined following the considerations and the notation of Curtice and Ettenberg [8].

(4) The dc state variables are $\mathbf{v} = [v_g \ v_d \ v_{dg}]^T$ for the Materka and Kacprzak model and $\mathbf{v} = [V_{in} \ V_{out} \ V_{dg}]^T$ for the Curtice and Ettenberg model.

TABLE III

UNITS OF THE FET MODEL PARAMETERS

Parameter	Unit	Parameter	Unit
R_g	Ω	I_s	A
R_i	Ω	α_s	1/V
R_d	Ω	I_{sr}	A
R_s	Ω	α_{sr}	1/V
G_{ds}	$1/\Omega$	I_{dss}	A
g_m	$1/\Omega$	α_d	-
L_g	nH	V_{po}	V
L_d	nH	γ	-
L_s	nH	C_{go}	pF
C_{dg}	pF	V_{bi}	V
C_{ds}	pF		
C_{gs}	pF		
τ	ps		

TABLE IV

PARAMETER VALUES OF THE FET MODEL

Parameter	Bias 1		Bias 2		Bias 3	
	start	solution	start	solution	start	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

Parameter	start	solution
I_{dss}	0.2	0.1888
α_d	4.0	3.0523
V_{po}	-4.0	-4.3453
γ	-0.2	-0.3958
C_{go}	1.0	0.6137
V_{bi}	1.0	1.3011

See bias conditions in Table V.

\+ values may not be reliable as the decomposition dictionary shows weak identifiability.

* values determined by $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and dc solution.

TABLE V

DC RESPONSES AND MEASUREMENTS

DC current	Bias 1	Bias 2	Bias 3
	V_{gs} = 0V V_{ds} = 4V	V_{gs} = -1.74V V_{ds} = 4V	V_{gs} = -3.10V V_{ds} = 4V
I_{gs} assumed	0.0A	0.0A	0.0A
i_{gs} calculated	-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} calculated	0.177A	0.092A	0.043A

TABLE VI

DECOMPOSITION DICTIONARY AT THE SOLUTION

Parameter	I_{gs}	I_{ds}	S_{11}	S_{21}	S_{12}	S_{22}
R_g	0.00	0.00	0.02	0.01	0.01	0.02
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_{go}	-	-	40.	26.	49.	9.9
V_{bi}	-	-	9.3	5.5	9.7	2.4

The dictionary was set up using 50 random points over a 25 percent range around the solution point.

TABLE VII

DECOMPOSITION DICTIONARY AT AN INTERMEDIATE STAGE

Parameter	S_{11}	S_{21}	S_{12}	S_{22}
R_g	1.7	0.56	1.2	0.16
L_g	1.1	0.27	0.84	0.08
L_d	0.05	0.19	0.65	1.2
R_d	0.07	0.70	1.1	2.9
R_i^1	0.48	0.16	0.98	0.05
R_i^2	0.41	0.14	0.51	0.06
R_i^3	0.41	0.10	0.39	0.05
I_s	0.00	0.00	0.00	0.00
α_s	0.00	0.00	0.00	0.00
I_{sr}	0.00	0.00	0.00	0.00
α_{sr}	0.00	0.00	0.00	0.00

Notes:

(1) Only relevant function groups and possible parameter candidates are listed.

(2) The dictionary is constructed by assuming an initial value of 0.01nH for L_g and L_d.

(3) 50 points and a 25 percent range were used to set up the dictionary.

Figure captions:

Fig. 1 Simple RC linear circuit example.

Fig. 2 (a) The Materka and Kacprzak nonlinear FET model [2], (b) the corresponding small-signal equivalent circuit.

Fig. 3 (a)The Curtice and Ettenberg nonlinear FET model [8], (b) the corresponding small-signal equivalent circuit.

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

Fig. 5 The S-parameter match at the solution of Case 2 using a simplified model for V_{ds} = 4V and V_{gs} = 0V.

Fig. 6 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.

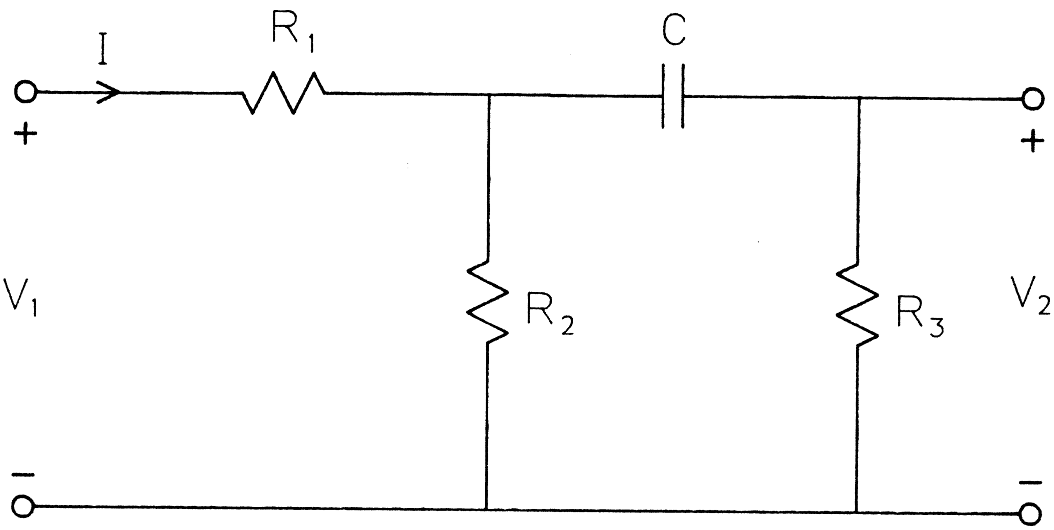


Fig. 1 Simple RC linear circuit example.

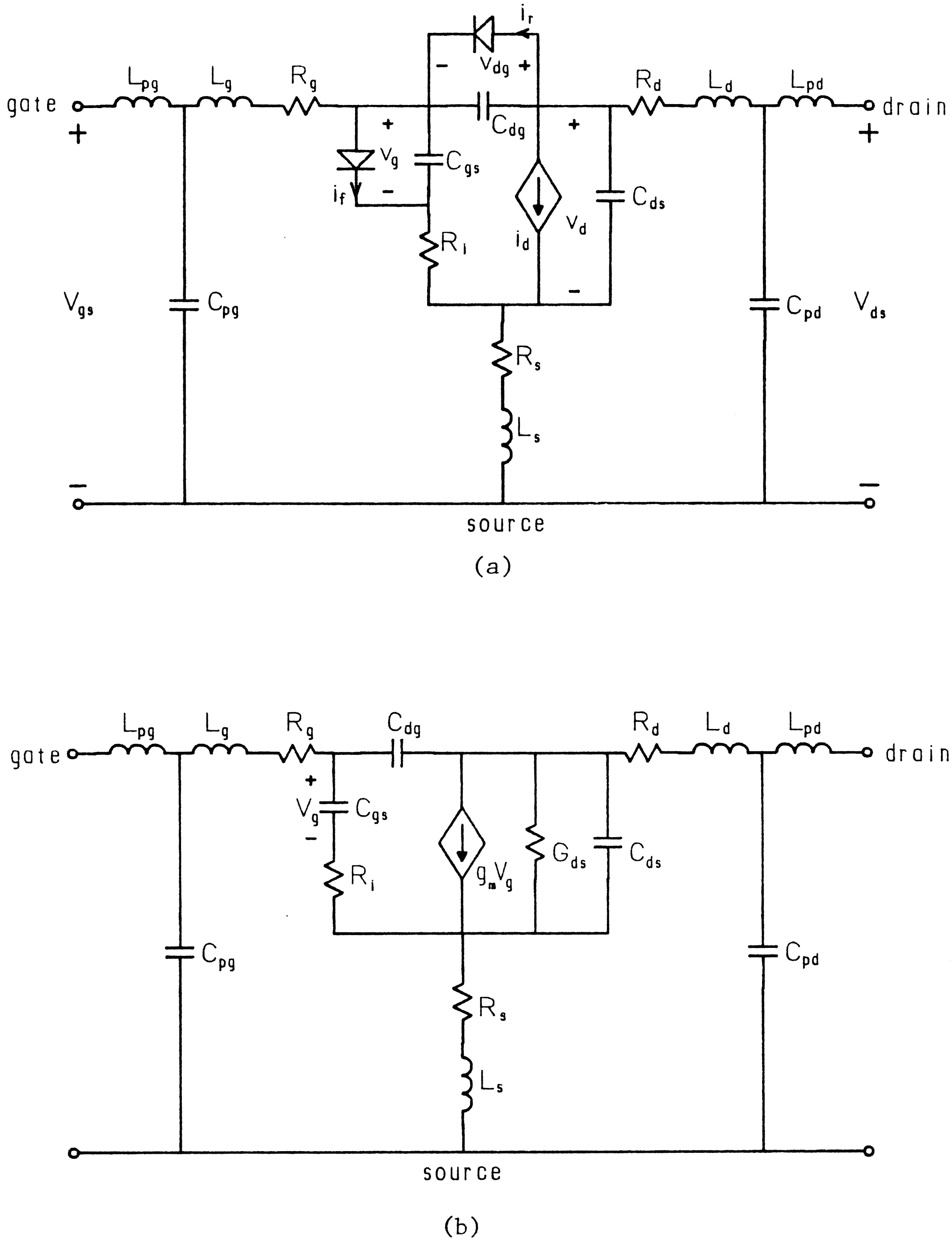
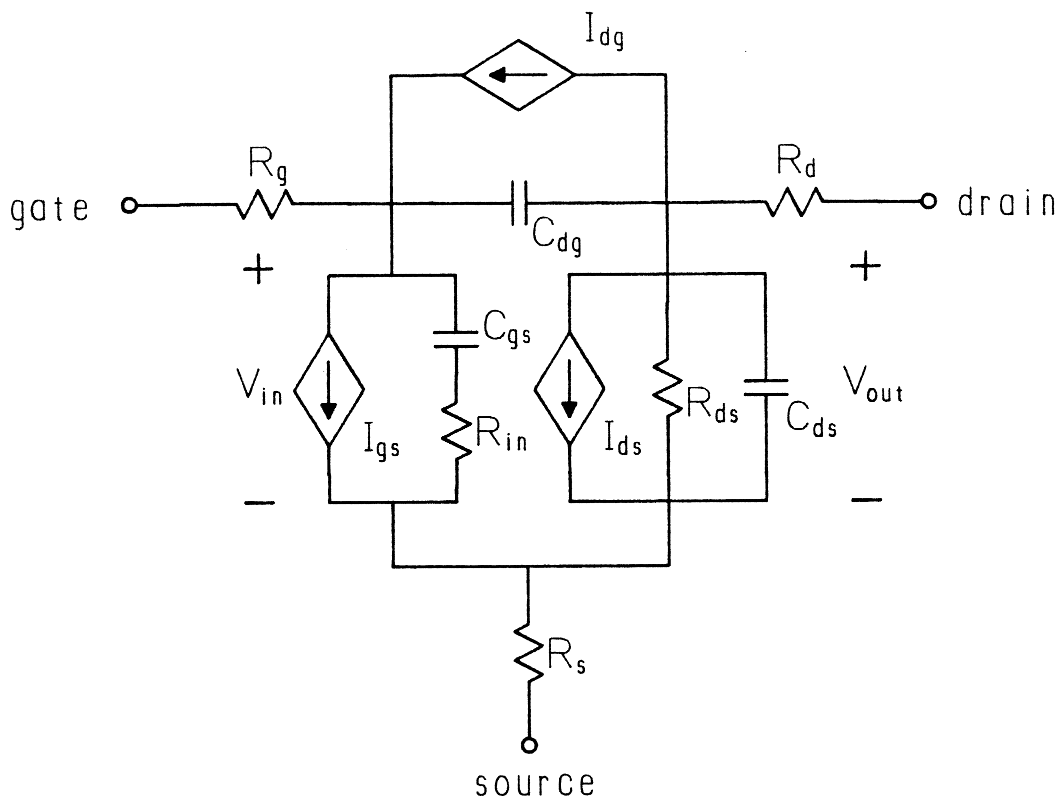
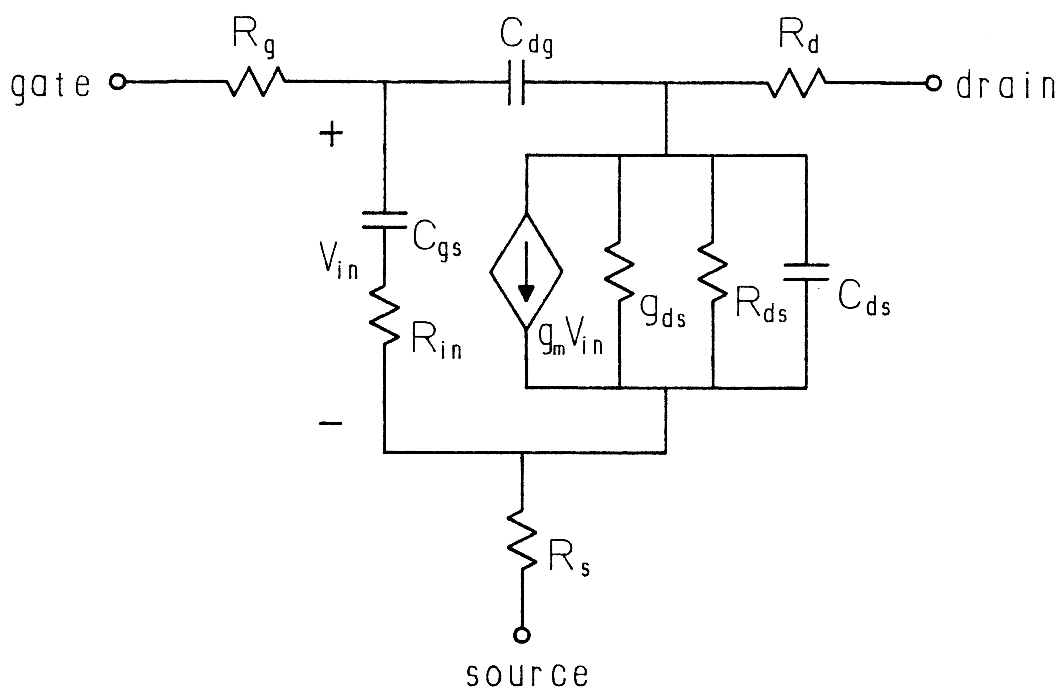


Fig. 2 (a) The Materka and Kacprzak nonlinear FET model [2], (b) the corresponding small-signal equivalent circuit.



(a)



(b)

Fig. 3 (a)The Curtice and Ettenberg nonlinear FET model [8], (b) the corresponding small-signal equivalent circuit.

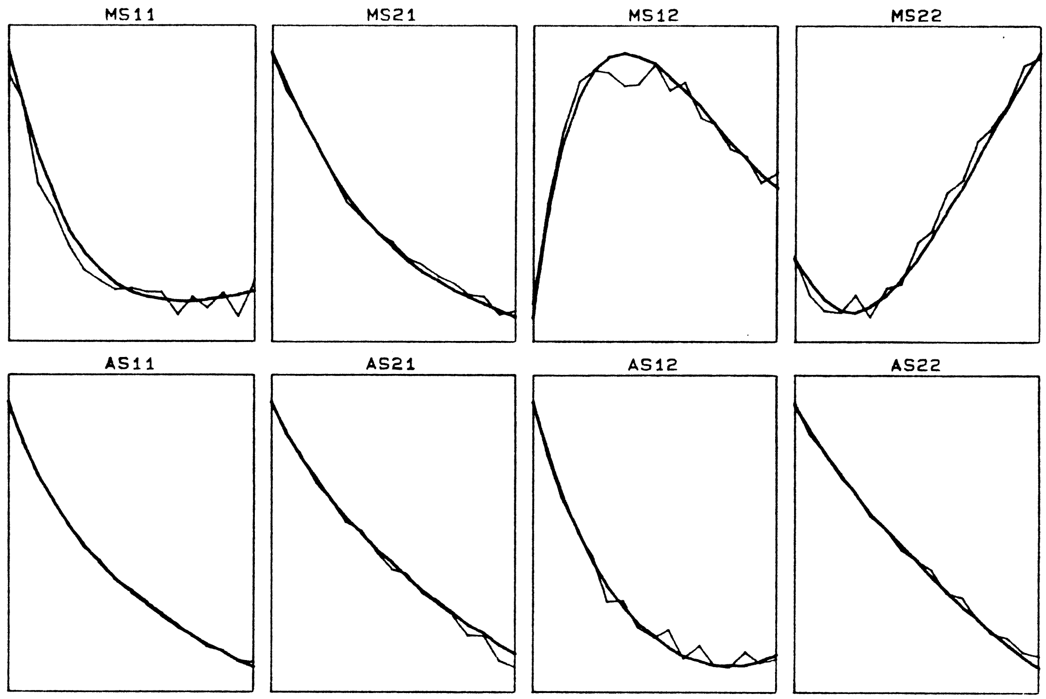


Fig. 4 The S-parameter match at the solution of Case 1 for V_{ds} = 4V and V_{gs} = 0V.

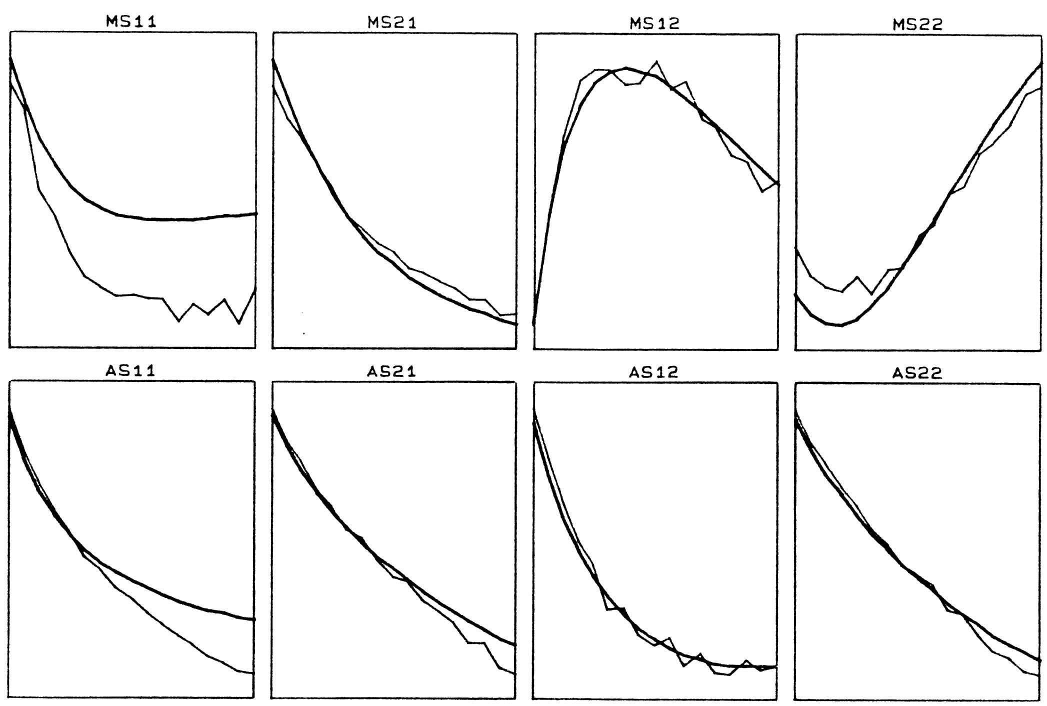


Fig. 5 The S-parameter match at the solution of Case 2 using a simplified model for V_{ds} = 4V and V_{gs} = 0V.

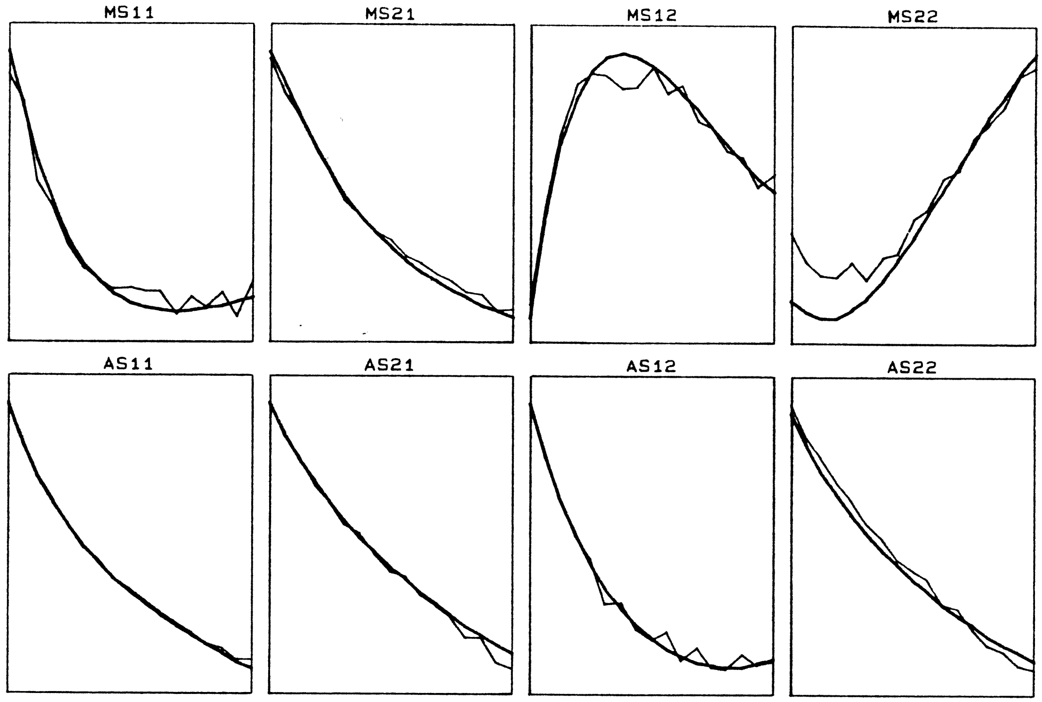


Fig. 6 The S-parameter match at the solution of Case 2 for V_{ds} = 4V and V_{gs} = 0V. R_g and L_g were included as optimization variables.