Identifying Model Parameters of Semiconductor Devices Using Optimization Techniques

Josef Dobeš[1], Martin Grábner[1] and Lubomír Hruškovič[2]

[1]Department of Radioelectronics, Czech Technical University in Prague, Czech Republic
[2]Department of Electromagnetic Field, Czech Technical University in Prague, Czech Republic

The optimization is an indispensable tool for extracting the parameters of any complicated models. Hence, advanced optimization techniques are also necessary for identifying the model parameters of semiconductor devices because their current models are very sophisticated (especially the BJT and MOSFET ones). The equations of such models contain typically one hundred parameters. Therefore, the measurement and particularly identification of the full set of the model parameters is very difficult. In the paper, an optimization method is presented which is applicable for the identifications of very complicated models using a relatively small number of iterations. The algorithm has been implemented into the original software tool C.I.A. (Circuit Interactive Analyzer) to its static and dynamic analysis modes. Therefore, the optimization is able to identify both direct-current and capacitance models of semiconductor devices. The process is demonstrated with various transistors.

Keywords: modeling, semiconductor device, measurement, parameter extraction, optimization, BJT, MOSFET, JFET

1. Introduction

For identifying the model parameters, both special and generally usable tools are used. The special tools usually work in multistep mode (i.e., the parameters or their groups are extracted successively), and therefore they are more robust than the generally usable ones. The extraction procedure for the EKV 2.6 MOSFET model [1] is the typical representative of this group. On the other hand, the MATLAB Optimization Toolbox [2] is one of the main representatives of the generally usable tools (this implementation contains several powerful optimization methods – e.g., preconditioned conjugate gradient,

Levenberg-Marquardt, or Gauss-Newton). The procedures of such tools usually determine the model parameters en bloc, and therefore they are not so reliable as the special tools. Moreover, they often need large number of iterations. However, they are able to optimize more complicated structures.

The C.I.A. optimization procedure belongs to the generally usable ones – the algorithm seeks to find up to 25 (in the current stable version of the program) unknown parameters of the circuit for the fulfillment of user-specified requirements. The algorithm controls the analyses and changes these parameters after each of them to successively fulfill the user's requirements.

2. Description of the Optimization Procedure of the C.I.A. Program

Let us assume that two circuit output variables are to be monitored at three points as shown in
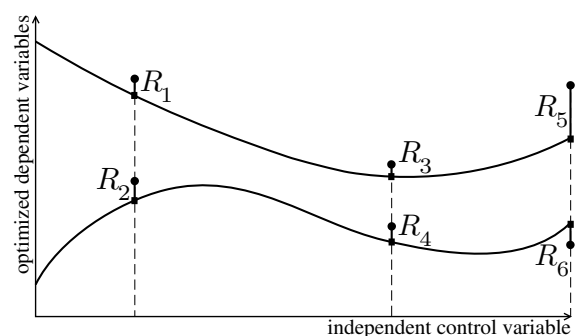

Figure 1. Diagram of a typical optimization task.

Figure 1. The circles mark the user-specified requirements on the output variables, and the squares mark values of the output variables obtained after an analysis. The algorithm seeks to minimize the sum of squares of differences between them

$$S(x_1, \ldots, x_n) = \sum_{k=1}^{m} R_k^2(x_1, \ldots, x_n), \ n \leqq m, \quad (1)$$

where the optimized parameters of the circuit are marked by $x_1, \ldots, x_n$, and $R_k, \ k = 1, \ldots, m$ are the differences.

An extreme (local or global) of the function of n variables (1) is found in the standard way, i.e., solving system

$$\nabla S = \sum_{k=1}^{m} 2R_k \nabla R_k = \mathbf{0}. \quad (2)$$

After the standard derivation (e.g., see [3]), the generalized least-squares procedure is obtained applying the condition (2) (^{t} marks matrix transposing)

$$J^t J \Delta x^{(l)} = -J^t r, \ x^{(l+1)} = x^{(l)} + \Delta x^{(l)}, \quad (3)$$
$$l = 1, \ldots, l_{\max},$$

where l is the iteration index and

$$r_k = R_k[x^{(l)}], \qquad \frac{\partial r_k}{\partial x_i} = \frac{\partial R_k}{\partial x_i}[x^{(l)}],$$

$$J = \begin{bmatrix} \dfrac{\partial r_1}{\partial x_1} & \cdots & \dfrac{\partial r_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial r_m}{\partial x_1} & \cdots & \dfrac{\partial r_m}{\partial x_n} \end{bmatrix}, \quad (4)$$
$$k = 1, \ldots m, \ i = 1, \ldots, n.$$

The generalized least-squares procedure is very fast, but sometimes insufficiently robust. Therefore, the method is combined with the classical gradient one

$$\Delta x^{(l)} = -2J^t r, \ l = 1, \ldots, l_{\max}$$

to the more reliable (Levenberg-Marquardt) modification of (3)

$$[J^t J + \lambda^{(l)} \mathbf{1}] \Delta x^{(l)} = -J^t r,$$
$$x^{(l+1)} = x^{(l)} + \Delta x^{(l)}, \ l = 1, \ldots, l_{\max}, \quad (5)$$

where $\mathbf{1}$ is unit matrix and $\lambda^{(l)}$ is a scalar iteration-dependent factor. There are many methodologies to optimally determine that factor at each iteration – the most sophisticated ones use an estimation based on the eigenvalues of the Jacobian (4) [4]. However, simpler empirical ways are mostly also successful [2,3]. The procedure of the C.I.A. program also contains an original version of the empirical method (however, the method based on the eigenvalues of $J^t J$ is also developed), which tries to decrease the $\lambda^{(l)}$ factor successively (i.e., to make the generalized least-squares method more influential at the end of the process):

$$\lambda^{(1)} = 1,$$
$$\lambda^{(l+1)} = \frac{\lambda^{(l)}}{5}. \quad (6)$$

However, this monotone decrease must be interrupted (and therefore the gradient method must be sometimes made more influential) when the method seems to diverge:

$$\texttt{if } l > 1 \wedge S^{(l)} \geqq \min_{j=1, \ldots, l-1} S^{(j)} \texttt{ then}$$
$$x^{(l)} := x^{(l-1)}, \ \lambda^{(l)} := \lambda^{(l)} 5^2, \quad (7)$$

where the first multiplication by 5 compensates the division by 5 in (6), and the second multiplication by 5 increases the scalar factor of the method.

Let us emphasize that the procedure (5)–(7) does not use usual one-dimensional minimizations. This is why the method of the empirical determination of the scalar factor is quite different from that in MATLAB [2]. However, the suggested procedure is appreciably faster – it needs tens or hundreds of iterations compared with thousands typically necessary for the same tasks in MATLAB (we tested three built-in methods).

Unfortunately, the method described above is still insufficient for a relatively wide class of the circuit optimization problems (especially for those which generate very different elements in the Jacobian regarding their magnitudes). Therefore, the following additional improvements were implemented in the C.I.A. procedure:

- the differences R_k in (4) must be normalized;
- the differences should also be weighted so that a measurement inaccuracy can be considered;
- the Jacobian $\boldsymbol{J}$ in (4) must *also* be normalized;
- the Jacobian should be determined quickly, using the sensitivity analysis;
- evaluating the Jacobian is not necessary in each iteration – the criterion has also been developed;
- a logarithmic damping suppressing possible divergence of iterations (5) has also been included.

2.1. Normalizing the System Equations

The models of semiconductor devices contain expressions with extreme differences of their magnitudes (tiny terms together with huge ones). For such systems, many of the standard optimization algorithms [3] are numerically unstable. Hence, a normalization of (5) is necessary. First, the differences are normalized together with their weighting:

$$R_k'[\boldsymbol{x}^{(l)}] \triangleq w_k \frac{y_k^{(\text{ident})}\left[\boldsymbol{x}^{(l)}\right] - y_k^{(\text{meas})}}{|y_k^{(\text{meas})}| + y_k^{(\text{null})}}, \qquad (8)$$
$$k = 1, \dots, m,$$

where $^{(\text{meas})}$ and $^{(\text{ident})}$ mark the measured and identified values, and the parameters $y_k^{(\text{null})}$ stabilize (8) when some measured values are near or equal to zero. However, many numerical experiments have proven that a normalization of the Jacobian is also necessary:

$$\frac{\partial R_k'\left[\boldsymbol{x}^{(l)}\right]}{\partial x_i} := w_k \frac{\partial y_k^{(\text{ident})}\left[\boldsymbol{x}^{(l)}\right]}{\partial x_i}$$
$$\times \frac{x_i^{(\text{max})} - x_i^{(\text{min})}}{|y_k^{(\text{meas})}| + y_k^{(\text{null})}}, \qquad (9)$$
$$k = 1, \dots, m, \ i = 1, \dots, n,$$

where $\frac{\partial y_k^{(\text{ident})}}{\partial x_i}$ is the output of the sensitivity analysis.

The equation (8) itself is a definition. However, the equation (9) represents an assignment

modified by the normalization. Therefore, the solution of the linear system in (5) must be modified by the assignment

$$\Delta x_i^{(l)} := \Delta x_i^{(l)}[x_i^{(\text{max})} - x_i^{(\text{min})}], \ i = 1, \dots, n$$

after each iteration, where $x_i^{(\text{min})}$ and $x_i^{(\text{max})}$ represent minimum and maximum allowable values specified by the user. These limits are mostly determined by the physics of semiconductor devices.

The optimization is one of the important advantages of the C.I.A. program compared with the other tools for CAD – it may be applied upon the operating-point, direct-current transfer, frequency, and transient analyses. The number of optimized circuit parameters is limited to 25. However, there is no problem to increase that number because the convergence does not depend on the task dimension. As the C.I.A. program is the generally usable tool, it is also possible to identify the parameters of composed structures, such as the Darlington couple or BJT-MOSFET cascode.

The empirical factor 5 in (6) and (7) has been carefully selected by means of many typical optimization tasks. It is the appropriate compromise between robustness and efficiency. For checking whether the found minimum of (1) is the global one, a semiautomatic method has been developed which uses automatically generated starting points. For semiconductor devices, this procedure is mostly sufficient (the problem of many local minima is more considerable for identifying the model parameters of transmission lines).

3. Results of the Model Identifications

The model equations which were used for the following identifications are defined in the appendix of [5]. A detailed physical theory on modeling the semiconductor devices is available in [6].

3.1. BJT

Low Frequency Transistor. The first identified BJT was KC508, which is a Czech equivalent of BC108. The transistor was firstly identified without the quasisaturation part of the model, which was simpler, of course. The results of the identification are shown in Figures 2 and 3 – the first one (forward mode) with the root mean square (rms) error 9.61 % and maximum absolute value of relative differences ($\delta_{\max}$) 43.1 %, and the second one (reverse mode) with the values rms = 4.85 % and $\delta_{\max} = 20.0$ %.

The optimization determined the values of the model parameters $I_S = 7 \times 10^{-13}$ A, $I_{SE} = 2.98 \times 10^{-11}$ A, $I_{SC} = 1.5 \times 10^{-11}$ A, $\beta_F = 974$, $\beta_R = 50$, $n_F = 1.1$, $n_R = 1.1$, $n_E = 2.06$, $n_C = 1.69$, $V_{AF} = 14.9$ V, $V_{AR} = 4.9$ V, $I_{KF} = 1.2$ A, $I_{KR} = 1.28$ mA, and $r_C = 3.2$ Ω.

As shown in Figure 2, the saturation part of the characteristics is not optimally modeled. Therefore, the newer part of the equations for modeling the quasisaturation [6] must also be considered. The results of such improved identification are shown in Figure 4 (they are drawn using natural linear axes for a comparison with the previous logarithmic ones). The optimization determined the additional model parameters $r_{CO} = 10$ Ω, $V_O = 100$ V, and $\gamma = 10^{-7}$. With the inclusion of the quasisaturation model, the errors of the identification were lesser than those above: rms = 3.51 % and $\delta_{\max} = 14.9$ %.
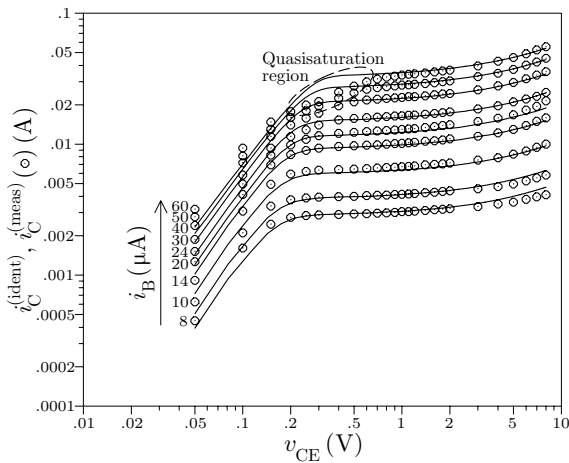

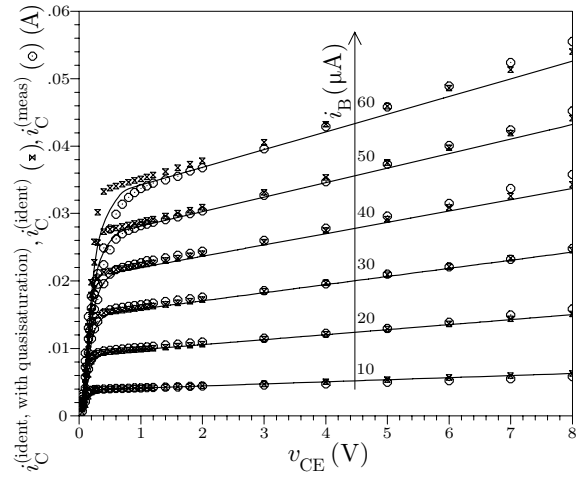*Figure 2.* Forward DC characteristics of the BJT KC508.

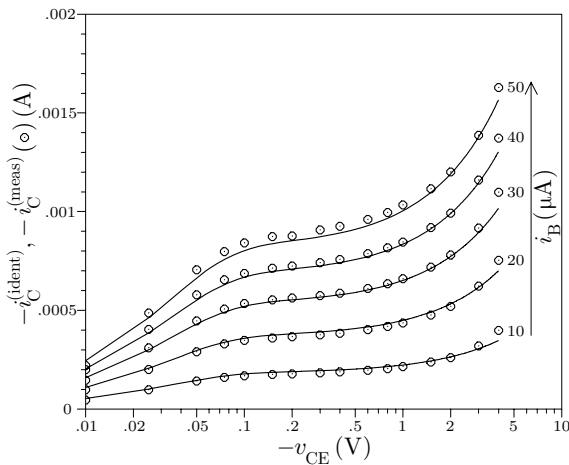*Figure 4.* Using quasisaturation model of the BJT KC508.

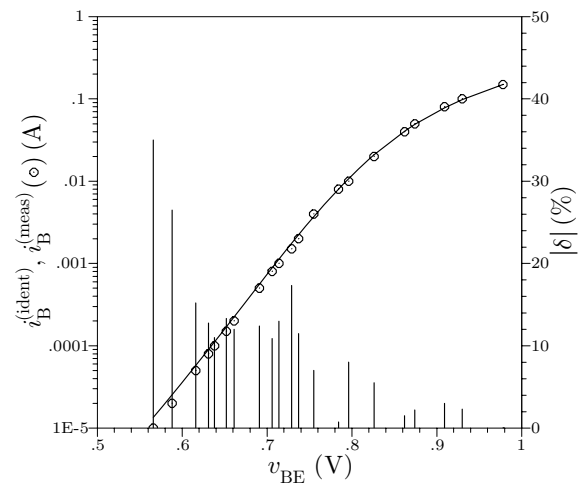*Figure 3.* Reverse DC characteristics of the BJT KC508.

Figure 5. Forward input characteristic of the BJT KC508.

The parameters of the nonlinear base-resistance model were identified using the input characteristic of the transistor as shown in Figure 5. The input characteristic was identified with the errors rms = 13.5 % and $\delta_{\max}$ = 35.0 % and the optimization determined the model parameters r_B = 26 Ω, r_{BM} = 37 mΩ, I_{rB} = 3.4 μA, and r_E = 0.53 Ω.

The parameters of the dynamic part of the model were also identified. First, both junction capacitances were determined as shown in Figure 6. The identification had the errors rms = 1.57 % (E), 1.64 % (C) and $\delta_{\max}$ = 2.51 % (E), 2.73 % (C), and the optimization gave the model parameters C_{JE} = 4.38 pF, ϕ_E = 0.65 V, m_E = 0.4, C_{JC} = 3.11 pF, ϕ_C = 0.4 V, and m_C = 0.273. Second, the transit-time model
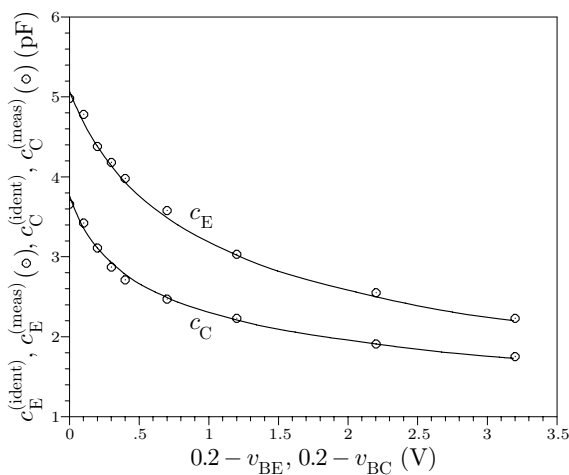
parameters were identified as shown in Figure 7. The optimization determined the model parameters τ_F = 0.249 ns, $I_{\tau F}$ = 0.35 A, $V_{\tau F}$ = 8.52 V, and $X_{\tau F}$ = 0.33 with the errors rms = 31.8 % and $\delta_{\max}$ = 94.4 %. The last two ones seem to be large – however, the differences were determined using the "vertical" distances which were not optimal here, of course (actually, the identification can be considered quite successful). The reverse transit time was identified in the same way with the result τ_R = 23 ns.

High Frequency Transistor. The second identified BJT was the microwave KT391 with the characteristics shown in Figure 8. The irregularities were probably caused by oscillations during the measurement – it was very difficult
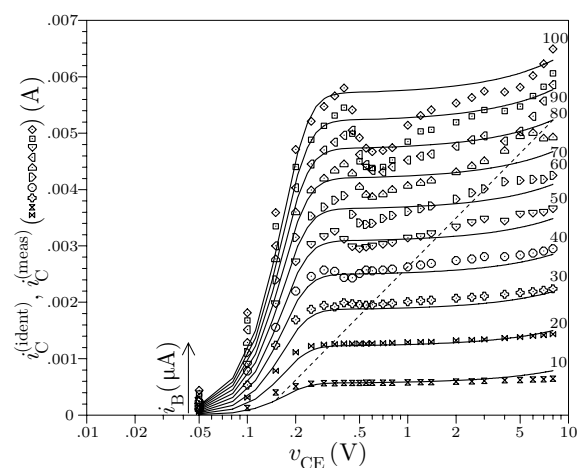

Figure 6. Collector and emitter junction capacitances of the BJT KC508.

Figure 8. Forward DC characteristics of the microwave BJT KT391.

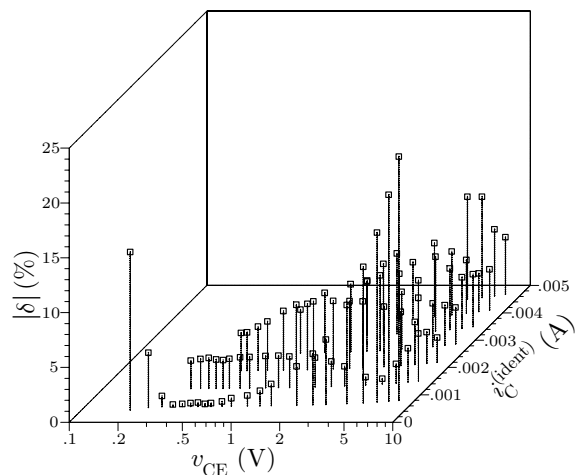*Figure 7.* Identification of transit-time model parameters of the BJT KC508.

Figure 9. Relative errors of the identification in the selected stable area.

to perform the DC measurements for the microwave transistors due to problematic stability.

The optimization determined the values of the model parameters $I_S = 10^{-8}$ A, $I_{SE} = 4.7 \times 10^{-9}$ A, $I_{SC} = 10^{-7}$ A, $\beta_F = 133$, $\beta_R = 1.6$, $n_F = 1.15$, $n_R = 1.13$, $n_E = 1.86$, $n_C = 1.75$, $V_{AF} = 123$ V, $V_{AR} = 2$ V, $I_{KF} = 18$ mA, $I_{KR} = 86$ mA, $r_C = 2\,\Omega$, $r_B = 10\,\Omega$, $r_{BM} = 1\,\Omega$, $I_{rB} = 100\,\mu$A, and $r_E = 1.6\,\Omega$ with the identification errors rms $= 16.0\,\%$ and $\delta_{\max} = 61.7\,\%$. However, if only the triangular "stable" region was used, as shown in Figures 8 and 9, then the identification errors were lesser: rms $= 5.99\,\%$ and $\delta_{\max} = 22.2\,\%$.

3.2. MOSFET

Enhancement Mode Transistors. At first, let us identify the models of enhancement-mode transistors. The first one was the low-power P-channel MOSFET 2N3608 – see Figure 10. The identification procedure determined the values of the model parameters $V_{TO} = -4.77$ V, $\phi_S = 0.657$ V, $\phi_O = 0.806$ V, $W = 37.9\,\mu$m, $L = 3.46\,\mu$m, $X_J = 1.54\,\mu$m, $X_{JL} = 0.762\,\mu$m, $t_{\text{ox}} = 98.7$ nm, $N_{FS} = 10^{15}$ m^{-2}, $N_A = 2.32 \times 10^{22}$ m^{-3}, $v_{\max} = 3.55 \times 10^5$ m/s, $\mu_O = 0.0719$ m$^2/$(Vs), $E_P = 3.4$ MV/m, $\kappa = 0.441$, $K_P = 2.49 \times 10^{-5}$ A/V^2, $\gamma = 0.294\,\sqrt{\text{V}}$, $\delta = 0.989$, $\eta = 0.03$, $\theta = 0.00334$ V^{-1}, and $\iota = 0.34$ (the last one was present only in the C.I.A. program where it served as an ad-
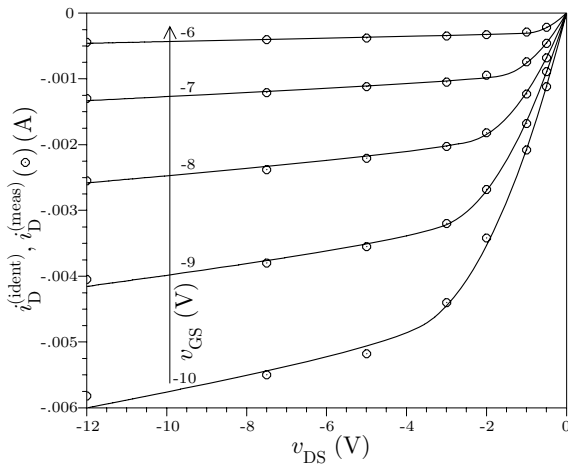

Figure 10. Forward DC characteristics of the P-channel enhancement-mode MOSFET 2N3608.
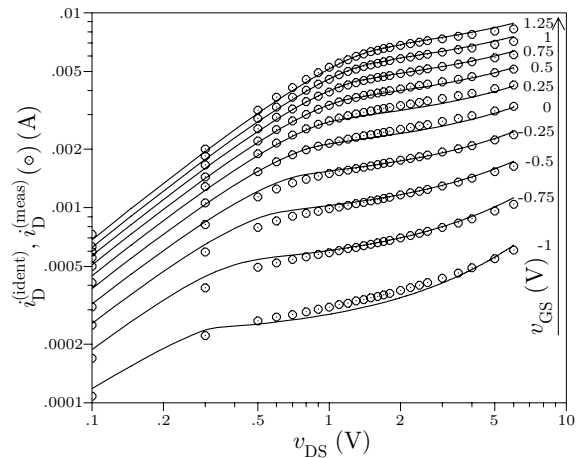

Figure 12. Forward DC characteristics of the N-channel depletion-mode MOSFET KF521.
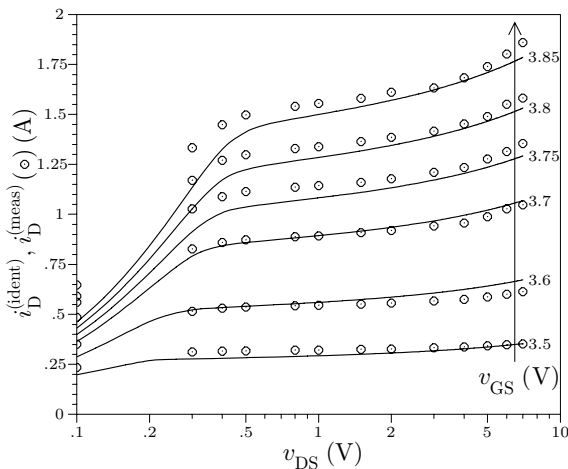

Figure 11. Forward DC characteristics of the N-channel enhancement-mode MOSFET BUZ345.
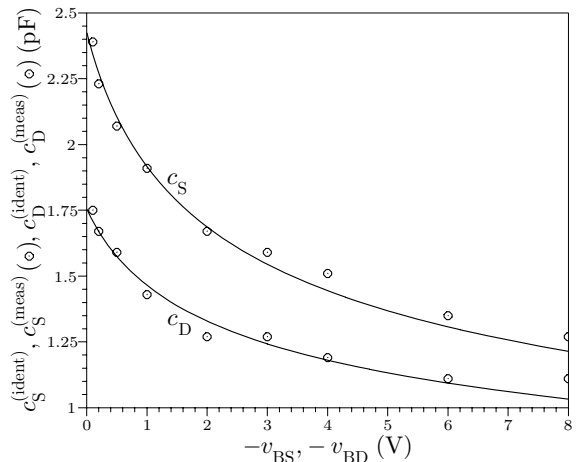

Figure 13. Drain and source junction capacitances of the MOSFET KF521.

ditional fitting factor). The parameters of the model were found with the excellent precision: rms = 2.18 % and δ_{max} = 5.41 %.

The second one was the high-power N-channel VMOS BUZ345 – see Figure 11. The identification procedure determined the values of the model parameters V_{TO} = 3.26 V, ϕ_S = 0.578 V, ϕ_O = 0.801 V, W = 1.46 m, L = 4.97 μm, X_J = 0.289 μm, X_{JL} = 0.179 μm, t_{ox} = 74.7 nm, N_{FS} = 10^{15} m^{-2}, N_A = 1.73×10^{20} m^{-3}, v_{max} = 3.23×10^5 m/s, μ_O = 0.0585 m^2/(Vs), κ = 0.0306, K_P = 4.19×10^{-5} A/V^2, γ = 0.366 $\sqrt{V}$, δ = 1, θ = 0.0384 V^{-1}, ι = 0.572, r_D = 0.0249 Ω, and r_S = 0.0435 Ω (for the power devices, the drain and source resistances must also be identified; in the previous example, their values were fixed to the defaults 10 Ω). The identification errors for that power device were greater than those for the previous one (which is natural): rms = 8.67 % and δ_{max} = 28.8 %. Moreover, the value of W was extreme but logical – power devices were composed of many single structures and therefore such value represented an integral.

Depletion Mode Transistor. Let us identify the model of a depletion-mode transistor which was the N-channel KF521 – see Figure 12. The identification procedure determined the values of the model parameters V_{TO} = -1.48 V, ϕ_S = 0.334 V, ϕ_O = 0.789 V, W = 443 μm, L = 4.83 μm, X_J = 0.932 μm, X_{JL} = 0.827 μm, t_{ox} = 71.8 nm, N_{FS} = 10^{15} m^{-2}, N_A = 7.51×10^{21} m^{-3}, v_{max} = 1.71×10^5 m/s, μ_O = 0.0535 m^2/(Vs), E_P = 419 kV/m, κ = 0.4, K_P = 2.12×10^{-5} A/V^2, γ = 0.568 $\sqrt{V}$, δ = 1, η = 0.811, θ = 0.002 V^{-1}, ι = 0.929, r_D = 11.8 Ω, and r_S = 5.17 Ω. Again, the identification ended with small errors rms = 4.06 % and δ_{max} = 14.5 %.

For the MOSFET KF521, the parameters of the model of its junction capacitances were also identified – see Figure 13. The identification procedure determined the model parameters C_{JO} area$_S$ = 2.17 pF, C_{JO} area$_D$ = 1.57 pF, $C_{JO\text{sw}}$ perimeter$_S$ = 0.26 pF, $C_{JO\text{sw}}$ perimeter$_D$ = 0.182 pF, ϕ_O = 0.789 V, $\phi_{O\text{sw}}$ = 0.789 V, m_S = 0.302, $m_{S\text{sw}}$ = 0.183, m_D = 0.213, and $m_{D\text{sw}}$ = 0.286 – again, the errors of the identification were relatively very small: rms =

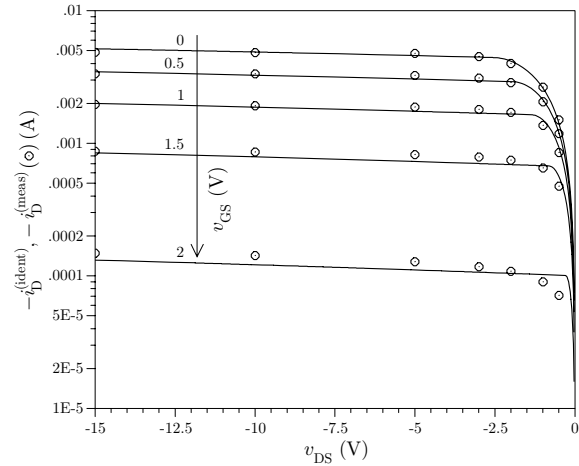*Figure 14.* Forward DC characteristics of the P-channel JFET 2N2498.

2.73 % (S), 3.15 % (D), and δ_{max} = 4.36 % (S), 6.90 % (D).

3.3. JFET

Let us identify the model parameters of the P-channel JFET 2N2498 – see Figure 14. The identification procedure determined the values of the model parameters V_T = -2.288 V (i.e., the "physical" threshold voltage was $+2.288$ V [6]), β = 1.299×10^{-3} A V^{-2}, λ = 0.02322 V^{-1}, r_D = 55.75 Ω, and r_S = 108.3 Ω with the errors rms = 10.4 % and δ_{max} = 42.38 % (the larger δ_{max} occurred for the voltages/currents near to zero, which is far from the standard JFET operating points).

4. Conclusion

An optimization algorithm has been presented which is convenient for the robust and efficient identifications of complicated models. The algorithm has been improved using the normalization of equations, which is important for stability of the identifications of semiconductor devices. The modified algorithm has been implemented into the C.I.A. program, and typical measurements and identifications of the model parameters have been demonstrated.

5. Acknowledgments

References

[1] M. BUCHER, C. LALLEMENT, C. C. ENZ, An Efficient Parameter Extraction Methodology for the EKV MOST Model. *In IEEE International Conference on Microelectronic Test Structures*, (1996).

[2] T. COLEMAN, Y. ZHANG, *Optimization Toolbox for Use with MATLAB$^®$*, User's Guide. The MathWorks, Inc., 2nd ed., 2003.

[3] R. FLETCHER, *Practical Methods of Optimization*. John Wiley & Sons, 1978.

[4] L. FINSCHI, An implementation of the Levenberg-Marquardt algorithm. Eidgenössische Technische Hochschule Zürich, 1996.

[5] A. VLADIMIRESCU, *The SPICE Book*. John Wiley & Sons, 1994.

[6] G. MASSOBRIO, P. ANTOGNETTI, *Semiconductor Devices Modeling With SPICE*. McGraw-Hill, 1993.

Contact addresses:
Josef Dobeš
Dept. of Radioelectronics
Czech Technical University in Prague
Technická 2, 166 27 Praha 6
Czech Republic
e-mail: `dobes@feld.cvut.cz`

Martin Grábner
Dept. of Radioelectronics
Czech Technical University in Prague
Technická 2, 166 27 Praha 6
Czech Republic
e-mail: `grabner@testcom.cz`

Lubomír Hruškovič
Dept. of Electromagnetic Field
Czech Technical University in Prague
Technická 2, 166 27 Praha 6
Czech Republic
e-mail: `hruskl2@feld.cvut.cz`

JOSEF DOBEŠ received the Ph.D. degree in microelectronics from the Czech Technical University in Prague in 1986. From 1986 to 1992, he was a researcher of TESLA Research Institute where he performed analyses on algorithms for CMOS Technology Simulators. He is currently with the Department of Radio Engineering of the Czech Technical University in Prague. His current research interests include physical modeling of elements of electronic circuits, especially radio-frequency and microwave transistors and transmission lines, creating or improving special algorithms for circuit analysis and optimization such as time- and frequency-domain sensitivity, poles-zeros or steady-state analyses, and creating a comprehensive CAD tool for analysis and optimization of radio-frequency circuits.

MARTIN GRÁBNER received the B.S. and M.S. degrees in electrical engineering from the Czech Technical University in Prague in 1998 and 2000, respectively. He has been with the Dept. of Microwave Communications, TESTCOM. His work is focused on the propagation of radio waves in the atmosphere. He is also interested in circuit modelling of microwave transistors and optimization techniques.

LUBOMÍR HRUŠKOVIČ was born in Levice, Slovak Republic, on May 30, 1973. In 2006, he received the Ing. degree from the CTU FEE, Prague, Czech Republic, where he is currently working toward his Ph.D. degree in the Department of Electromagnetic Field.