DOI: https://doi.org/10.30898/1684-1719.2025.10.5

EXTRACTION OF MOSFET MODEL PARAMETERS IN UNIFORM APPROXIMATION

V.N. Biryukov

Institute for Radio Engineering Systems and Control of Southern Federal University, 347900, Russia, Rostov Region, Taganrog, Nekrasovsky, 44

The paper was received June 25, 2025.

Abstract. To the best our knowledge, this work is the first attempt of extraction of MOSFET parameters in the uniform (Chebyshev) approximation. The proposed optimization algorithm uses random descent with a random number generator whose probability density is best suited for the ravine structure of the objective function in such a way as to obtain a compromise between the probability of a successful attempt and the magnitude of the descent step. The proposed distribution reduces the descent time by two orders of magnitude compared to the uniform one. By replacing the root-mean-square approximation with a uniform one, the maximum error of the model can be significantly reduced, which improves the accuracy of the worst-case analysis of electronic circuits.

Key words: extraction of MOSFET parameters, least squares, random optimization, uniform approximation.

Corresponding author: Biryukov Vadim Nikolaevich, vnbiryukov@yandex.ru

Introduction

The last stage of the analysis of electronic devices in practice is always the worst-case analysis [1]. The key role in such an analysis is played by the correspondence of transistor models and experimental characteristics. To analyze such a relationship, a statistical analysis of the parameters of the required type of transistor is carried out, as a result of which the manufacturing variations in parameters is defined, allowing the analysis of the circuit by the Monte Carlo method.

Extraction of transistor parameters is currently carried out, as a rule, by the least squares method, which guarantees a minimum mean square error of the model, but does not guarantee a minimum of maximum error. With parameters obtained by the least squares method, the maximum error of the model exceeds the root mean square by up to 3 times [2], that is, the estimate of the actual maximum error in this case turns out to be greatly overestimated. The actual maximum model error in the uniform approximation is not always slightly larger than the mean square error. There is always a possibility that the model whose parameters are extracted by least squares has a maximum error twice as large as the model whose parameters are determined by optimization in the uniform approximation.

There are three reasons why uniform approximation is not used in parametric optimization in practice. First, replacing the «small» mean square error with a «large» maximum error is “unpleasant” [3]. Second, the results obtained with uniform approximation are much more sensitive to the error and the number of measurements of the device characteristics than those obtained by the least squares method. Third, and most importantly, in uniform approximation the objective function is non-differentiable, and the zero-order methods used, which do not determine the derivatives of the objective function, are extremely slow: the number of evaluations of the objective function can reach 10^{11} [4].

Manufacturing parameter deviations are measured on a sample of hundreds of devices, i.e. the parameter extraction procedure is repeated many times. In addition, manufacturing sometimes involves continuous testing, so fast gradient least squares optimization methods are used despite their high maximum error.

When optimizing for any criterion, it is necessary to take into account that the objective function usually has a ravine structure and is called a valley function [5]. Then in any method the descent from the starting point initially occurs quickly along the steep sides of the ravine, and then passes slowly along the gently sloping bottom. In such a problem, small changes in the objective function at the bottom of the ravine correspond to large changes in some parameters, which lead to uncertainty in the choice of the stopping criterion. In this case, the error estimate of nonlinear least squares remains insufficiently reliable [6], which is also true for uniform approximation.

An empirical method for accelerating random descent for problems with a ravine structure of the objective function is known, which consists in decreasing the interval of change of random parameters as the descent proceeds [7]. In this paper, it is proposed to use a special law of distribution of random parameters for the same purpose, which does not require the selection of empirical constants dependent on the type of ravine structure and the initial conditions of the descent.

1. Random optimization algorithm

The basic random optimization sequential algorithm for any objective function $Q(\mathrm{x})$ can be described as:

Initialize x with a random position in the search-space.

● Until a number of iterations performed, repeat the following:

●● Sample a new position y by adding a random vector to the current position x.

●● If $Q(\mathrm{y}) < Q(\mathrm{x})$ then move to the new position by setting x = y.

●Now x holds the best-found position.

In the simplest case, each element of the vector y_k is selected on the interval $(x_k - \Delta_k/2,\ x_k + \Delta_k/2)$, where Δ_k is the width of the interval, and on this interval y_k is a random variable with some distribution. Conflicting requirements are imposed on the width of the search interval. On the one hand, to increase the descent step, interval must be increased, and to increase the probability of successful y_k, in the bottom of ravine it must be reduced. This contradiction is most acute in problems of high dimension of the objective function and/or poor conditionality of the Hessian matrix

2. Selection of the distribution law of random parameters

In Fig. 1, the polygons show the level surfaces of the two-dimensional ravine objective function and the squares in which the trial parameters are randomly assigned during the descent.

The probability of a successful attempt is determined by the ratio of the area of the figure, limited by the perimeter of the square and the part of the square located below the level passing through the center of the square to the area of the square. When descending the slope of a ravine, regardless of the size of the square, this probability is constant and equal to 0.5. In this case, it is obviously advantageous to increase the size of the square, since this increases the step of the descent.

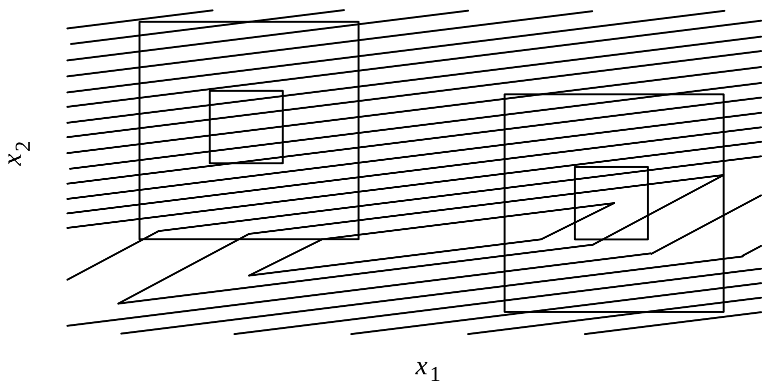


Fig. 1. Level surfaces of a two-dimensional ravine objective function in uniform approximation. The squares correspond to possible regions of random parameters during descent: on the left at the slope of the ravine, and on the right at the bottom of the ravine.

At the bottom of the ravine, as follows from Fig. 1, the probability of a successful attempt depending on the size of the square has a more complex nature. If the size of the square is reduced to less than the size of the side of the polygon of the nearest level surface, then the probability of a successful attempt changes slightly, but the solution step decreases. If the side of the square is increased above the same side of the polygon, then the probability of a successful attempt drops sharply. Thus, at each point on the bottom of the ravine, the size of the square is optimal if it coincides with the nearest smallest side of the level surface, that is, as you move along the bottom, the square should decrease.

In work [7], it is proposed to reduce the region of admissible values of random parameters during the descent according to some empirical law, but the efficiency of such a method depends on both the type of the objective function and the initial approximation.

The above contradiction can be partially removed by choosing a rapidly changing distribution function of the random variable y_k, that has maximum near x_k and minimum at the boundaries of the interval Δ_k. Here y_k is random variable with the uniform cubic distribution $f_3(x) = 1/(6x^{2/3})$ [8]. Each value of the uniform cubic distribution corresponds to a cube of uniform distribution in the interval [–1, 1]. Fig 2 shows the main feature of probability density function of cubic uniform distribution as $\lim_{x \to 0} f_3(x) = \infty.$ This property allows for efficient random descent both down a slope and along the entire bottom of a ravine. A normally distributed random vector gives much less efficiency.

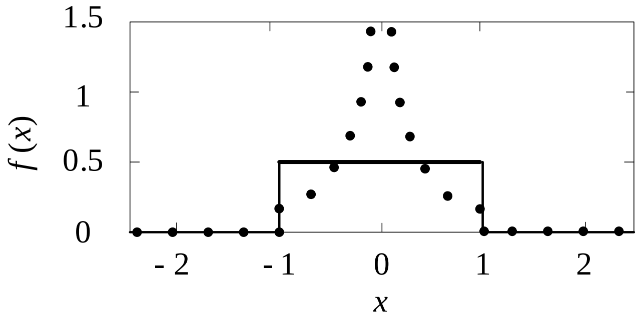


Fig. 2. Probability density function. The solid and dotted lines show the uniform distribution and cubic uniform distribution, respectively.

3. MOSFET model

To demonstrate the properties of the random descent method, the MOSFET model from [9] was chosen as

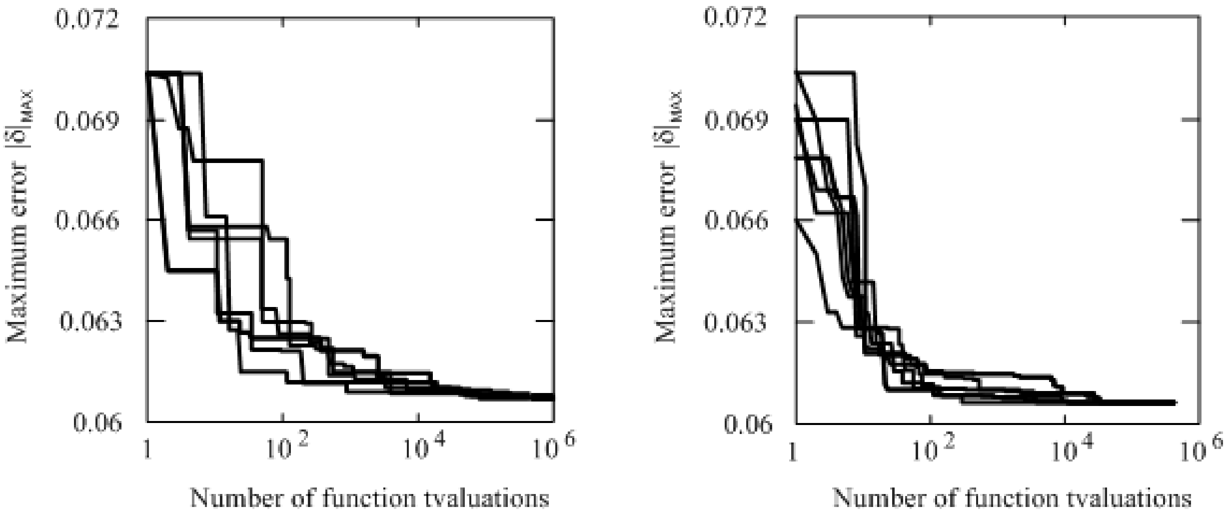
$$I_D = \begin{cases} 0; & V_{GS} \le V_{TH}; \\ \beta V_K \left(V_{DS} - V_K \ln \dfrac{V_K + V_{GS} - V_{TH}}{V_K + V_{GS} - V_{TH} - V_{DS}} \right); & V_{DS} < V_{ST}; \\ \beta V_K \left(V_{ST} - V_K \ln \dfrac{V_K + V_{GS} - V_{TH}}{V_K + V_{GS} - V_{TH} - V_{ST}} \right) & \\ \times \dfrac{1 + a(V_{DS} - V_{ST})}{1 + b(V_{DS} - V_{ST})}; & V_{DS} \ge V_{ST}, \end{cases} \tag{1}$$

where I_D is a drain current; V_{GS} and V_{DS} are gate-source, drain-source voltages, respectively; V_{TH} denotes a threshold voltage; β is an intrinsic transconductance parameter; V_K is transverse electric field parameter; $V_{ST} = k(V_{GS} - V_{TH})$ is a drain saturation voltage; k is empirical constant ($0 < k < 1$). Constants a and b are expressed explicitly from the conditions of continuity of the current and its first and second derivatives at $V_{DS} = V_{ST}$ as

$$b = -\frac{1}{2} \left. \frac{\partial^2 I_D}{\partial V_{DS}^2} \right|_{V_{ST}} \Bigg/ \left. \frac{\partial I_D}{\partial V_{DS}} \right|_{V_{ST}}; \; a = b + \left. \frac{\partial I_D}{\partial V_{DS}} \right|_{V_{ST}} \Bigg/ I_D(V_{GS}, V_{ST}).$$

4. Random descent with different distributions of random parameters

Fig. 3 shows the sequence of maximum error of model (1) along the descent for uniform and cubic uniform distribution of parameters and $\Delta_k = 0.05\, x_k$ for GaN MOSFET GS66508B [10].



a) b)

Fig. 3. Maximum error versus the number of function evaluations. (a) Uniform distribution. (b) Cubic uniform distribution.

Since the graphs are random, their families are shown to improve reliability. The initial values of the parameters were obtained by least squares with a root mean square error of 2.5 %. With uniform approximation, the RMS error increases to 3.5 %. The transition to a cubic uniform distribution reduces the number of objective function estimates in the case under consideration by two orders of magnitude.

Note that the cubic uniform distribution accelerates the descent significantly only at the bottom of the ravine. Thus, this distribution allows us to obtain not so much a more accurate value of the minimum of the objective function, but to refine the determined parameters of the model and make the criterion for stopping the descent more reliable.

Above, we used the L_p-norm for $p = 2$ for initial guess. Using the L_p-norm for $p > 2$ allows us to obtain initials at lower values of $|\delta|_{MAX}$. However, as p increases, the conditionality of the Hessian matrix decreases, so the maximum achievable value of p is low [11]. For various models and transistors, it has been experimentally established that $4 \leq p_{MAX} \leq 32$. With increasing p, $|\delta|_{MAX}$ decreases monotonically, but some model parameters decrease non-monotonically, so a limited increase in p has little effect on the duration of random descent.

Conclusion

The proposed method allows obtaining the exact minimum maximum error of the transistor model is simple and does not require complex subroutines. We expect that if there is a solution for the least squares method, then it also exists for the uniform approximation. The method does not introduce additional empirical coefficients, and can be used with any other methods of accelerating random descent.

References

1) Lenertz B.A., Command A.F.S. Electrical design worst-case circuit analysis: Guidelines and draft standard (REV A) // TOR-2013–00297, The Aerospace Corporation. – 2013. – P. 109.
2) Pilipenko A.M., Biryukov V.N., Prokopenko N.N. A template model of junction field-effect transistors for a wide temperature range // 2019 IEEE East-West Design & Test Symposium (EWDTS). – IEEE, 2019. – P. 1-4. http://doi.org/10.1109/EWDTS.2019.8884411
3) Hand M.L. Aspects of linear regression estimation under the criterion of minimizing the maximum absolute residual. – Iowa State University, 1978.
4) Gorbunov E., Dvurechensky P., Gasnikov A. An accelerated method for derivative-free smooth stochastic convex optimization // SIAM Journal on Optimization. – 2022. – T. 32. – №. 2. – P. 1210-1238. https://doi.org/10.1137/19M1259225
5) Rakitskii Yu.V. Minimization methods for functions depending strongly on a few variables. Encyclopedia of Mathematics. [Online]. Available: https://encyclopediaofmath.org/wiki/Minimization_methods_for_functions_depending_strongly_on_a_few_variables
6) Dennis Jr J.E., Schnabel R.B. Numerical methods for unconstrained optimization and nonlinear equations. – Society for Industrial and Applied Mathematics, 1996. https://doi.org/10.1137/1.9781611971200
7) Luus R., Jaakola T.H.I. Optimization by direct search and systematic reduction of the size of search region // AIChE Journal. – 1973. – T. 19. – №. 4. – P. 760-766. https://doi.org/10.1002/aic.690190413

8) Pilipenko A.M., Biryukov V.N. Efficiency improvement of the random search algorithm for parametric identification of electronic components models // 2016 International Siberian Conference on Control and Communications (SIBCON). – IEEE, 2016. – P. 1-6. http://doi.org/10.1109/SIBCON.2016.7491703
9) Biryukov V.N. Models of the wide-bandgap MOSFET // Journal of Radio Electronics. – 2024. – №. 10. https://doi.org/10.30898/1684-1719.2024.10.17
10) Gansystems Datasheet GS66508B-DS [Online] Available: https://gansystems.com/wp-content/uploads/2019/05/GS66508B-DS-Rev-190502.pdf
11) Bandler J.W., Charalambous C. Practical least pth optimization of networks // IEEE Transactions on Microwave Theory and Techniques. – 1972. – T. 20. – №. 12. – P. 834-840. https://doi.org/10.1109/TMTT.1972.1127894

For citation:

Biryukov V.N. Extraction of transistor model parameters in uniform approximation // Journal of Radio Electronics. – 2025. – №. 10. https://doi.org/10.30898/1684-1719.2025.10.5