

MOSFET parameter extraction with EKV model

Lukasz Bartnik and Jarosław Arabas

Warsaw University of Technology, Institute of Electronic Systems, Warsaw, Poland,
email: L.Bartnik, jarabas@elka.pw.edu.pl

Abstract

This paper reports results of an ongoing research on MOSFET parameter extraction using the EKV model. This work continues efforts in finding the best method for efficient and robust parameter extraction based on voltage-current structure characteristics. In the extraction process, voltage-current characteristics are matched by the characteristics generated by the model. Values of parameters for which the best match is observed are the result of the extraction. The extraction process is considered an optimization problem, which is then solved by an evolutionary algorithm followed by the Nelder-Mead simplex. The influence of the measurement error on the extraction results is investigated experimentally.

1 Introduction

This paper reports an ongoing research on the MOSFET (Metal-Oxide Semiconductor Field-Effect Transistor) parameter extraction. CMOS (Complementary Metal-Oxide Semiconductor) is currently the leading technology in the field of integrated circuits (IC) production. More details on the background of this work can be found in [2, 3] where the first part of this research has been presented.

The main goal of our work is to develop a method for extracting values of a set of transistor parameters, which would be computationally efficient and robust against measurement error. Our approach consists in comparing two sets of voltage-current I-V characteristics – (1) a set of output current values measured for the given input voltage values for a real MOSFET structure, and (2) a set of model-generated output current values for the same input voltages. When the model values match real ones for the same inputs, then the model parameters under extraction are assumed to match the respective values of the real structure.

The first attempt described in [2] was to use the Pierret-Shields model [7]. This is a relatively simple model simulating a number of physical effects present in the modeled class of transistors. A combination of an evolutionary algorithm and the Nelder-Mead simplex [6] was used to minimize the mean square error of matching measured and model-generated current-voltage characteristics. In this work the Pierret-Shields model has been replaced with the EKV model [4, 5], which is well recognized in the field of MOSFET modelling and takes into account a significantly larger number of physical MOSFET phenomena. The aim of experiments described in this paper was to evaluate the usability of the EKV model in the task of transistor parameter extraction.

We tried to determine the number of global or sufficiently good local optima for the objective function which is based on the EKV model. If there were more than one such optimum for a given set of I-V characteristics, then the extraction results would be ambiguous. The second aim of our work was to determine whether the measurement errors have any impact on the extraction results – and if true how strong it is.

The paper is organized as follows. Section 2 provides more details about the EKV model itself, the parameters under extraction, and the optimization methods. In section 3 the results of an experiment determining the number of global optima are described. Section 4 presents the results of an experiment considering the impact of the measurement error on the final extraction results. The last section draws some final conclusion and presents possible outlook.

2 Extraction method details

The extraction process described in this paper consists of two main components: the EKV transistor model and the optimization method. The latter uses two optimization algorithms: an evolutionary algorithm (AE) which is used to determine rough approximation of the solution, and is followed by the Nelder-Mead simplex, whose task is to perform fine tuning of the final solution. This approach is quite robust against getting stuck into poor local optima, and also shows good exploitation capabilities.

2.1 EKV model

In this work the EKV model version 2.6 was used. It was developed by Matthias Bucher of the Swiss Federal Institute of Technology [1], and was implemented in the C language. The model has 26 configurable electro-physical parameters, most of which have been set to constant values. Six of these parameters have been extracted, namely: *vto* (long-channel threshold voltage), *gamma* (body effect factor), *phi* (bulk Fermi potential), *kp* (transconductance parameter), *theta* (mobility reduction coefficient), and *ucrit* (longitudinal critical field).

Every parameter was constrained by its minimal and maximal value; this results in a set of box constraints, as following: *vto*: (0.0, 2.0), *gamma*: (0.0, 2.0), *phi*: (0.0, 2.0), *kp*: (0.0, 0.001), *theta*: (0.0, 0.2), and *ucrit*: (10^6 , 10^8).

To assure that every parameter is treated with the same importance in the optimization process, all of them have been scaled to (0.0, 1.0) range. Since the minimum and maximum values of parameter *gamma* are of two orders of magnitude different, the logarithm of its value is extracted.

2.2 Objective function

Adoption of the EKV model and the optimization algorithms in order to compose a parameter extraction method is as follows. Given are I-V characteristics measured for the MOSFET structure under extraction. Next, given a parameter vector \mathbf{p} and the input voltage values (V) from the I-V measurements, an output set is generated using the EKV model. Then, the model-generated output set and measured current values set (I) are compared for all input values and the mean square error (*mse*) is computed,

which is the optimizers' objective function to be minimized:

$$mse(\mathbf{p}) = \frac{1}{n} \cdot \sum_{i=1}^n (c_i - ekv(\mathbf{p}, \mathbf{v}_i))^2 \quad (1)$$

where n is the number of measurement points, \mathbf{v}_i is the i -th input voltage, c_i is the corresponding measured output current, and ekv denotes calculations performed by the EKV model.

2.3 Optimization methods

During all experiments described in this paper the same configuration for optimization methods was used. The evolutionary algorithm's population size was set to 15 individuals, the algorithm was stopped after reaching the number of 250 iterations, and the selection was non-elitist, fitness-proportional. The first population was initialized randomly with uniform distribution within the box constraints. Every individual was a subject to an uncorrelated Gaussian mutation with zero mean and standard deviation equal 0.02 – i.e. 2% of the whole value range for a single parameter. No crossover was used.

The Nelder-Mead simplex was initialized with the best parameter vector found by the evolutionary algorithm. It was stopped when the minimum distance between any pair of simplex vertices reached the value of $\sqrt{6} \cdot 10^{-5}$, which is 10^{-5} times the diagonal of the admissible set.

3 Determining the reliability of extraction results

The first two experiments were aimed at recognizing the objective function landscape, and especially at determining the number of good local optima. If there was any local optimum with the objective function value comparable to the global optimum, the extraction results would be ambiguous.

The first experiment, described in the following subsection, was designed to determine whether a known set of parameter values can be extracted from an artificially generated measurements set.

3.1 Re-discovering parameters of an existing structure

Values of parameters used to generate the artificial measurements set corresponded to an existing MOSFET structure and were as follows [4]. The six parameter under extraction were: $vto = 0.647$, $gamma = 0.78$, $phi = 0.93$, $kp = 4.304 \cdot 10^{-5}$, $theta = 0.026$, and $ucrit = 4 \cdot 10^6$. The rest of the parameters, which were set to constant values and were not a subject of extraction, were: $cox = 0.7 \cdot 10^{-3}$, $xj = 1.5 \cdot 10^{-7}$, $dw = -2.0 \cdot 10^{-8}$, $dl = -5.0 \cdot 10^{-8}$, $e0 = 0.0$, $tox = 6.5 \cdot 10^{-8}$, $lambda = 0.23$, $weta = 0.05$, $leta = 0.28$, $q0 = 2.8 \cdot 10^{-4}$, $lk = 5.0 \cdot 10^{-7}$, $iba = 2.0 \cdot 10^8$, $ibb = 3.5 \cdot 10^8$, $ibn = 1.0$, $tcv = 1.5 \cdot 10^{-3}$, $bex = -1.5$, $ucex = 1.7$, $ibbt = 0.0$, $avto = 1.0 \cdot 10^{-8}$, $akp = 2.5 \cdot 10^{-8}$, $agamma = 1.0 \cdot 10^{-8}$, $kf = 1.0 \cdot 10^{-27}$, $af = 1.0$, $xqc = 0.0$. Parameters $nsub$, vfb , $u0$, nqs , and $satlim$ were turned off – i.e. set to *NaN* value.

The set of measurement points was generated in the following way. A sequence of values distributed evenly within the admissible voltage range was used for every input voltage. This resulted in a set of input values distributed on a rectangular grid. Then,

for every input vector the output current was computed using the EKV model set up with the aforementioned parameters set. A number of different input voltage sets was generated, each with different grid granularity, which resulted in different set sizes.

Since the extraction method is non-deterministic, it was run 500 times, and the results were compared to see whether it is able to re-discover the same set of parameter values in each run. This was performed independently for every input set.

In Fig. 1 *mse* error plots are presented for each of the optimized parameters. Every point in the picture represents the value of a single parameter and the *mse* error value for a single final solution yielded by the hybrid of the evolutionary algorithm followed by the Nelder-Mead simplex. This means that with quite high probability each of these points is also a local optimum or is in a local optimum's vicinity.

The narrow end of each of tornado-like shapes visible in Fig. 1 matches the original parameter value used to generate the input set. Also, given that in each plot there is only one such narrow end, this may suggest that there is only one global minimum of the *mse* error function, and that it represents the original parameter vector used for input set generation. From plots in Fig. 1 the following conclusions can be also drawn.

First of all, the global minimum which is located in the lowest point of the tornado-like shape on each plot is rather hard to find. The hybrid method ends in local optima in most of its runs. Histogram of the logarithm of the *mse* error presented in Fig. 2 shows that there are at least two well-distinguishable groups of the extraction method result. One of them, with *mse* error logarithm greater or equal to -16 , comprises the vast majority of all results. The second group, with *mse* error logarithm value lower than -20 , is formed by the results located in the narrow end of the tornado-like shapes presented in Fig. 1.

Secondly, it becomes clear that since the extraction method has a non-deterministic character, only a fraction of final results is close enough to the actual parameter value to be considered reliable. Thus, only after a decent number of the extraction method runs, the best solution will be located in proximity of the global optimum. In Fig. 2 this fraction of best results comprises approximately 10% of all results.

3.2 Random structures re-discovering

In the second experiment the same procedure was performed, but this time for a hundred different random parameters sets. Every set was chosen with a uniform distribution within the box constraints. Then, for every set 500 independent runs of the extraction method were performed.

Table 1 presents mean value of the *mse* error and mean values of the extraction errors for each parameter. Parameter extraction error values are given in percent. All values are presented for percentiles: 50th, 25th, 15th, 5th, and 1st (best solution found). It is important to mention, that ordering by the misfit of characteristics measured by the *mse* error as defined in equation 1 might be different from ordering by the extraction error value specific to any of the parameters.

This experiment once again shows that due to the non-determinism of the extraction method it should be run many times to yield a satisfactory low level of the *mse* error. However, if the number of runs is large enough, the best obtained results are very close to the actual values for each parameter.

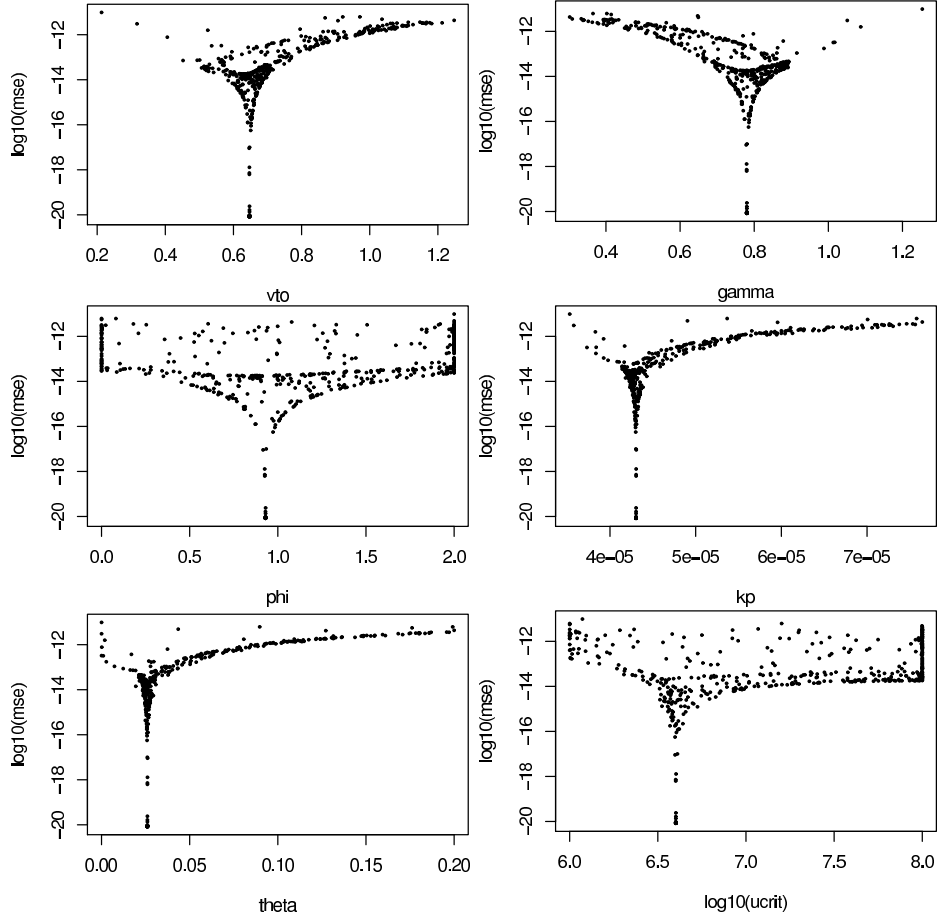


Figure 1. Plot of the logarithm of the matching error (Y axis) vs. each parameter value (X axis)

4 Impact of the measurement error on the extraction quality

In reality, every measurement is made with an error. We examine the influence of a non-systematic error assuming that the measured value x and its measurement \hat{x} are related according to:

$$\hat{x} = x + \varepsilon \quad (2)$$

where the random error is denoted by $\varepsilon \sim N(0, \sigma(x))$, and $\sigma(x)$ means that the standard deviation is a function of the measured value.

Two error values of the standard deviation were considered, namely 0.1% and 0.5% of the input value. For each error standard deviation the whole following procedure was performed.

Firstly, a set of parameter values to be re-discovered using the extraction method was chosen – these values were the same as presented in section 3.1. Then, artificial

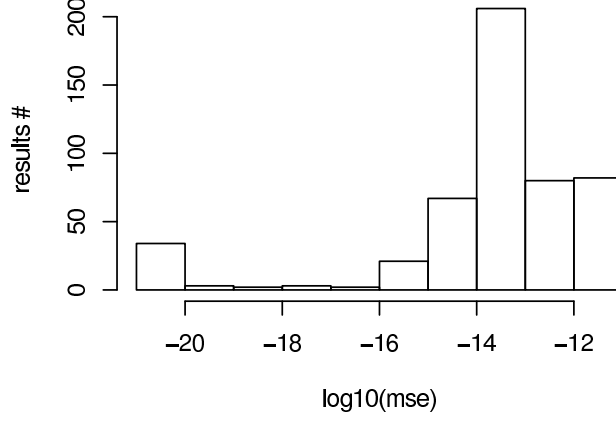


Figure 2. Histogram of the logarithm of the *mse* error

Percentile	<i>mse</i>	Mean relative error for					
		<i>vto</i>	<i>gamma</i>	<i>phi</i>	<i>kp</i>	<i>theta</i>	<i>ucrit</i>
50th	$7.817 \cdot 10^{-13}$	4.889	4.994	60.51	0.9203	6.294	6.051
25th	$4.442 \cdot 10^{-13}$	1.625	1.953	14.60	0.2752	2.882	2.372
15th	$1.045 \cdot 10^{-13}$	0.9608	1.556	8.438	0.1412	1.595	0.9124
5th	$5.172 \cdot 10^{-13}$	0.1165	0.94	7.048	0.0301	0.2658	0.2309
best	$1.74 \cdot 10^{-13}$	0.0014	0.171	0.195	0.0015	0.0083	0.0145

Table 1. Mean relative error for chosen percentiles – *mse* and for each parameter

measurement sets were generated according to the following procedure:

1. a set of input voltage values was generated,
2. every input voltage value was disturbed according to:

$$\hat{v}_{i,j} = v_{i,j} + \varepsilon \quad (3)$$

$$\varepsilon \sim N(0, \sigma(v_{i,j}))$$

where $v_{i,j}$ is the original input value, and $\sigma(v_{i,j})$ is the error's standard deviation,

3. output current c_i was computed using the EKV model and the disturbed input voltage values,
4. the output current value was then randomly disturbed with the standard deviation $\sigma(c_i)$ according to:

$$\hat{c}_i = ekv(\hat{\mathbf{v}}_i) + \varepsilon \quad (4)$$

$$\varepsilon \sim N(0, \sigma(ekv(\hat{\mathbf{v}}_i)))$$

where $\hat{\mathbf{v}}_i$ is the input disturbed by the random error, \hat{c}_i is the corresponding current-output measurement calculated for $\hat{\mathbf{v}}_i$ by the use of *ekv* denoting the EKV model,

5. the final artificial measurement set was obtained by combining the disturbed values from points 1. and 4.

Then, for each artificial measurements set 500 independent runs of the extraction method were performed. Tables 2 and 3 present the experiment results. From results gathered for each input set – with respect to the *mse* formula given in equation 1 – the best 5% was chosen. Then, for this top 5% of results, the average relative extraction error value for each parameter was computed. The error is expressed as the percentage of the original parameter value.

Input set size	Mean relative error for					
	<i>vto</i>	<i>gamma</i>	<i>phi</i>	<i>kp</i>	<i>theta</i>	<i>ucrit</i>
2197	0.125	0.220	1.330	0.022	0.506	0.470
729	0.382	0.228	1.976	0.008	0.052	0.093
343	1.018	0.336	4.957	0.072	1.253	0.207
125	0.005	0.720	2.457	0.302	5.454	1.701

Table 2. Average extraction error for top 5% of results, $\sigma = 0.1\%$

Input set size	Mean relative error for					
	<i>vto</i>	<i>gamma</i>	<i>phi</i>	<i>kp</i>	<i>theta</i>	<i>ucrit</i>
2197	1.97	4.47	28.12	1.09	9.55	0.14
729	3.0	6.8	46.0	1.6	10.0	2.9
343	0.47	2.23	5.09	0.63	12.83	3.11
125	6.2	3.0	29.4	1.1	7.1	1.0

Table 3. Average extraction error for top 5% of results, $\sigma = 0.5\%$

It can be seen that for $\sigma = 0.1\%$ the extraction error values are in most cases no greater than 1%. Apart from *phi* and *theta*, the extraction errors for the parameters are quite similar. Results obtained for $\sigma = 0.5\%$ presented in table 3 are visibly worse, which is a clear demonstration that the extraction quality goes down when the measurement error increases. For both values of the measurement error standard deviation, *phi* and *theta* seem to be the hardest to extract, while *vto*, *kp* and *ucrit* give most satisfactory error results.

No clear dependence of the extraction error on the number of measurement points can be observed. This may be due to the non-deterministic extraction method character.

In Fig. 3 *mse* error plots for each extracted parameter are presented. After introducing the measurement error tornado-like shapes around the original parameters values cannot be observed. This means that the fraction of results close to the global optimum, i.e. the original parameters set, cannot be distinguished under the *mse* criterion so easily this time.

It can be observed that the impact of the measurement error on the extraction results cannot be ignored. This is visible both in plots in Fig. 3 and in tables 2 and 3 containing the error values.

Even though the relative extraction error values seem to be quite significant, in fact error values no greater than $\pm 5\%$ are usually acceptable as the extraction results in the field of MOSFET parameter extraction.

Also, as mentioned before, error values are not directly correlated to the input set size, which might be due to the randomness in the extraction process. We suppose that with

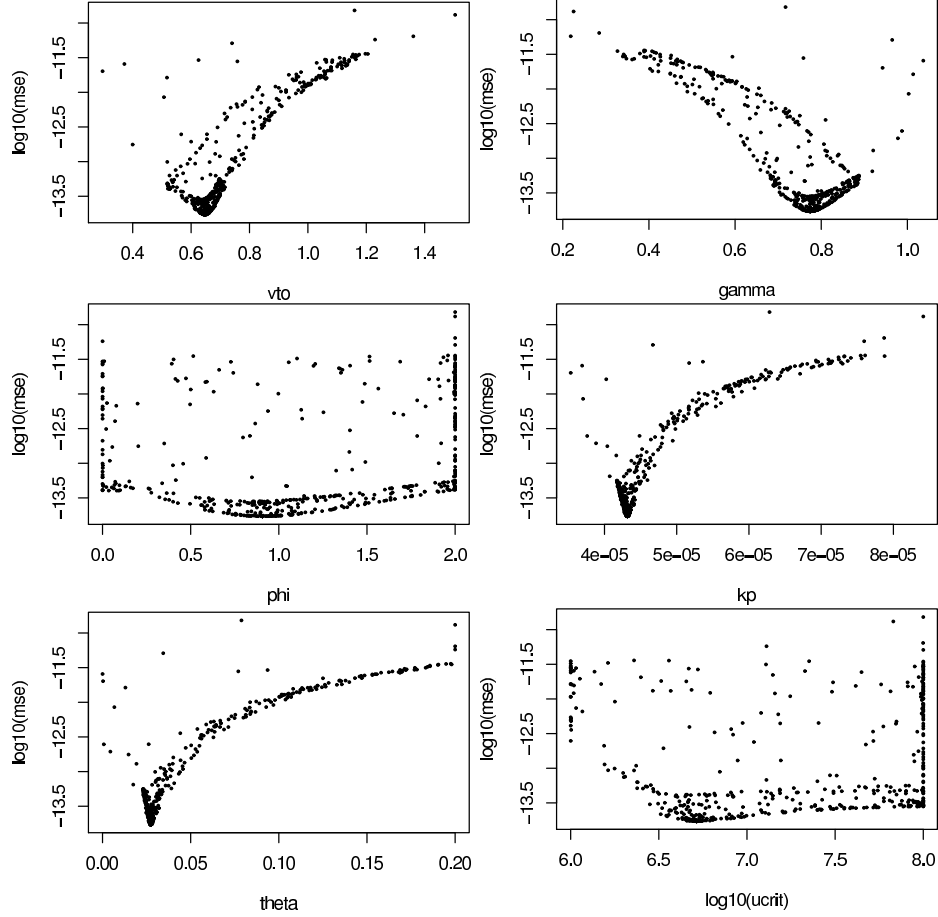


Figure 3. Plot of the logarithm of the matching error (Y axis) vs. each parameter value (X axis), measurement error $\sigma = 0.1\%$

bigger number of independent runs the correlation between the set size and extraction error would be higher. However, the relative differences of the extraction error among all parameters seem to be quite similar regardless of the presence or strength of the measurement error.

5 Conclusions and outlook

This paper presents a next step of research on robust and efficient extraction method for MOS transistor parameters. We have analysed the possibility of using the EKV model as the component of this method. First of all, it has been experimentally confirmed that a function returning the discrepancy between a real voltage-current measurement set and a EKV-generated one has one global optimum. This optimum can be found with a given level of probability. This allowed us to find out how the measurement error affects the

final extraction result. The conclusion is that, apart from the two worst parameters ϕ and θ , this method yields acceptable results.

We plan to gain a better insight on the measurement error propagation and possible methods which minimize its impact. Secondly, we believe that the extraction method can be handled in a number of steps, in each step a different subsets of parameters might be extracted from data. Finally, some field-specific extraction methods might be used to initialize the EA in the proximity of the actual global optimum.

6 Acknowledgements

We would like to express our gratitude to Daniel Tomaszewski, Institute of Electron Technology, Poland and Sławomir Szostak, Institute of Microelectronics and Optoelectronics, Warsaw University of Technology for their most appreciated help and advice during our work with the EKV model and interpreting the results.

We would also like to express our gratitude to Matthias Bucher and Władysław Grabiński, Swiss Federal Institute of Technology for sharing the EKV model implementation.

Bibliography

- [1] EKV Authors webpage. <http://legwww.epfl.ch/ekv/people.html>.
- [2] J. Arabas and S. Szostak. Hybrid Evolutionary Algorithm in MOSFET parameter extraction. *Prace Naukowe Politechniki Warszawskiej, Elektronika*, 2008.
- [3] J. Arabas, S. Szostak, L. Lukasiak, and A. Jakubowski. Studies of the feasibility of using global and local optimization methods in MOSFET characterization. *Elektronika*, 1, 2008.
- [4] M. Bucher, C. Lallement, F. Theodoloz, C. Enz, and F. Krummenacher. The EPFL-EKV MOSFET model equations for simulation, version 2.6. 1997. http://legwww.epfl.ch/ekv/pdf/ekv_v262.pdf.
- [5] C. Enz, F. Krummenacher, and E. Vittoz. An analytical MOS transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications. *Journal on Analog Integrated Circuits and Signal Processing*, pages 83–114, 1995.
- [6] J.A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, pages 308–313, 1965.
- [7] R.F. Pierret and J.A. Shields. Simplified long-channel MOSFET theory. *Solid-State Electron*, 26, 1983.

