MODELING AND SIMULATION OF TRANSISTOR MOSFET (HIGH-K) USING NEURAL NETWORK.

1Abdelkrim MOSTEFAI  2Smail BERRAH
1Applied Materials Laboratory, University of Sidi Bel Abbes, 22000- Sidi Bel Abbes, Algeria, mostakrimo@yahoo.fr.
2LMER Laboratory, University of A/ Mira of Bejaia, Algeria, sm_berrah@yahoo.fr.

3Hamza ABID
3Applied Materials Laboratory, University of Sidi Bel Abbes, 22000- Sidi Bel Abbes, Algeria, abid_hamza@yahoo.fr.

Abstract: Smart tools are increasingly used in the design, modeling and control of complex systems, in this paper; we present the results of the characterization and modeling of the electrical current-voltage of the MOSFET (HfO$_2$ dielectric oxide), using neural network.

Key words: Microelectronic, MOSFET, CMOS technology, high-k dielectric, artificial intelligence, neural network.

1. Introduction.

The reduction of the MOS transistors dimensions is accompanied by the reduction of the gate oxide thickness. In this context, one solution is to replace SiO$_2$ as a gate dielectric by high-dielectric constant (high-k) materials (for example HfO$_2$ oxide [1,2]) to reduce leakage currents.

Assuming that the MOSFET strongly submicron is a complex electronic device, electrical and physical phenomena (degradation, the effect of miniaturization, quantum effect, ...) Nonlinear, presenting difficulties of its study and its modeling, it can be studied using statistical methods in the field of artificial intelligence.

The paper addresses a simple and fast approach to implement Artificial Neural Networks (ANN) models for the MOS transistor (HfO$_2$ dielectric oxide) into MATLAB.

2. Continuous model ‘EKV MODEL’

The operation of a MOSFET can be separated into three different modes:

When $V_{GS} < V_T$ where $V_T$ is the threshold voltage of the device, the transistor is turned off, and there is no conduction between drain and source ($I_D \approx 0$).

When $V_{GS} > V_T$ and $V_{DS} \leq V_{GS} - V_T$. The transistor is turned on; the current from drain to source is modeled as:

$$I_D = \frac{\mu_n E_{ox} W}{2t_{ox}} \left[ (V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right]$$

Or

$$I_D = k_n [2(V_{GS} - V_T)V_{DS} - V_{DS}^2]$$

$$k_n = \frac{\mu_n E_{ox} W}{2t_{ox} L} = \frac{\mu_n C_{ox} (W/L)}{2}$$  \hspace{1cm} (3)

When $V_{GS} > V_T$ and $V_{DS} \geq V_{GS} - V_T$. The switch is turned on, and a channel has been created, which allows current to flow between the drain and source. When the device is in saturation the drain current is given by:

$$I_D = k_n (V_{GS} - V_T)^2$$  \hspace{1cm} (4)

When $V_{DS} = V_{GS} - V_T$

Into either Equation (2) or (4), so we get:

$$I_D = k_n V_{DS}^2$$  \hspace{1cm} (5)

The fig.1 presents the log $I_D$–$V_{GS}$ characteristics of a standard MOSFET.

![Fig.1. Discontinuity of the $I_{DS}$ ($V_{GS}$) characteristics at $V_{GS} = V_T$.](image)

In some cases, such as for the simulation of analog circuit, it is necessary to have a model which is continuous in all regime of operation (weak or strong inversion). In 1995, Enz, Krummenacher and Vittoz proposed a relatively simple MOSFET model
valid in all regions of operation: weak, moderate, and strong inversion. This has come to be known as the EKV model [3].

Their basic equation for drain current (in saturation) is given by [4,5]:

\[
I_D = I_F - I_R
\]

\[
I_D = 2n\mu_n C_{ox} \frac{W}{L} \left( \frac{K}{q} \right)^2 \left[ \ln \left( \frac{V_p - V_S}{2KT/q} \right) + \exp \left( \frac{V_p - V_S}{2KT/q} \right) \right]^2 - \left( \ln \left( 1 + \exp \left( \frac{V_p - V_{DS}}{2KT/q} \right) \right) \right)^2
\]

\[\quad (7)\]

\[
V_p = \frac{(V_{GS} - V_T)}{n}
\]

On the other hand, \(V_S = 0\), \(V_{DS} < V_p\) and \(V_{GS} > V_T\) (i.e. the transistor is operating in the non-saturated regime). In that case the exponential terms are much larger than unity, and one can write:

\[
I_D = 2n\mu_n C_{ox} \frac{W}{L} \left( \frac{K}{q} \right)^2 \left( \frac{V_p}{2KT/q} \right)^2 - \left( \frac{V_p - V_{DS}}{2VT/q} \right)^2
\]

\[\quad (9)\]

\[
I_D = \frac{1}{2} n\mu_n C_{ox} \frac{W}{L} \left( 2V_{DS}V_p - V_{DS}^2 \right)
\]

\[\quad (10)\]

\[
I_D = \frac{1}{2} n\mu_n C_{ox} \frac{W}{L} \left( \frac{2(V_{GS} - V_T)V_{DS}}{n} - V_{DS}^2 \right)
\]

\[\quad (11)\]

\[
I_D = \mu_n C_{ox} \frac{W}{L} \left( V_{GS} - V_T \right)V_{DS} - \frac{1}{2} nV_{DS}^2
\]

\[\quad (12)\]

3. Equivalent oxide thickness.

The miniaturization of the transistors results in the decrease of the gate length. The capacitive coupling between the gate and the substrate must increase or remain constant. Therefore, to maintain these properties, the gate oxide thickness \(t_{ox}\) is reduced [6], but when the thickness decreases below a certain value [7,8], several problems arise [1-9]. To resolve this problem, one should consider replacing the gate dielectric (currently \(SiO_2\)) by a higher permittivity dielectric (for example \(HfO_2\)), one can then increase the thickness and thereby reduce the gate current, while keeping a high \(C_{ox}\). Then introduce the notion equivalent oxide thickness (EOT):

\[
C_{SiO_2} = C_{High-k} \iff \frac{\varepsilon_{SiO_2}}{t_{SiO_2}} = \frac{\varepsilon_{High-k}}{t_{High-k}}
\]

\[\quad (13)\]

\[
\Rightarrow t_{SiO_2} = EOT = \frac{\varepsilon_{SiO_2}}{\varepsilon_{High-k}} t_{High-k}
\]

\[\quad (14)\]

\[
EOT = \frac{\varepsilon_{ox}}{\varepsilon_{High-k}} t_{High-k} + \frac{\varepsilon_{ox}}{\varepsilon_{SiO_2}} t_{SiO_2}
\]

\[\quad (15)\]

\[
C_{tot} = \frac{\varepsilon_{ox}E_0}{EOT}
\]

\[\quad (16)\]

The fig.2 presents the Comparison of leakage current density as a function of the equivalent oxide thickness between \(HfO_2\) and \(SiO_2\).

The fig.3 presents the Illustration of the integration of dielectric high-k in a MOSFET structure.
4. Neural network.

Multilayer Perceptron (MLP) is the most used neural network model. MLP utilizes a learning algorithm (backpropagation) for training the network [11]. MLP are capable to separate data that are not linearly separable [12]. The MLP structure is shown in Fig. 4.

5. Neural network transistor model.

The neural network designed to connect the input vector \((V_d, V_g, t_{ox}, \text{and } L)\) for MOSFET (HfO\(_2\) monolayer) (Fig. 5-a) and \((V_d, V_g, t_{ox}, t_{SiO2}, \text{and } L)\) for MOSFET (HfO\(_2\)/SiO\(_2\) multilayer) (Fig. 5-b) to output vector \(I_d\). Each of these parameters is indexed by a neuron (Fig. 5). The output activation function is the linear function, for hidden layers we choose the sigmoid function (logsig). To enhance learning, we modified different parameters such as the number of neurons in the hidden layer.

Table 1 summarizes the characteristics of the network optimized of the MOSFET.

<table>
<thead>
<tr>
<th>Propriety</th>
<th>Characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>10-8-1 MLP</td>
</tr>
<tr>
<td>Function of activation</td>
<td>Logsig-Logsig-linear</td>
</tr>
<tr>
<td>Learning rule</td>
<td>back propagation</td>
</tr>
<tr>
<td>EQM of test</td>
<td>0.0001</td>
</tr>
<tr>
<td>Itération</td>
<td>4000</td>
</tr>
</tbody>
</table>
The program of learning is launched. One observes then at the same times the evolution of the error to each epoch (cycle of learning), then the function obtained by the network, the evolution of the average error of learning is shown in Fig. 6.

Fig. 6. Evolution of the average error of learning.

6. Results and discussion.
Figures 7, 8, 9 and 10 show a comparison between the results predicted by the neuronal model (ANN) of different characteristics I-V (I_D-V_D and I_G-V_G) with those calculated by the analytical model for a transistor MOSFET (monolayer HfO_2) low doped with L=0.1 μm, t_{ox}=14 nm (monoclinic HfO_2, EOT=3 nm) and t_{ox}=22 nm (tetragonal HfO_2, EOT=3 nm).

Fig. 7. I_D (V_D) Characteristic of the MOSFET high-k monolayer HfO_2 (monoclinic-HfO_2 oxide).

Fig. 8. I_G (V_G) Characteristic of the MOSFET high-k monolayer HfO_2 (monoclinic-HfO_2 oxide).

Fig. 9. I_D (V_D) Characteristic of the MOSFET high-k monolayer HfO_2 (tetragonal-HfO_2 oxide).

Fig. 10. I_G (V_G) Characteristic of the MOSFET high-k monolayer HfO_2 (tetragonal-HfO_2 oxide).

Figures 11, 12, 13 and 14 show a comparison between the results predicted by the neuronal model (ANN) of different characteristics I-V (I_D-V_D and I_G-V_G) with those calculated by the analytical model for a transistor MOSFET (multilayer HfO_2/SiO_2) low doped with L=0.1 μm, t_{sio2}=1 nm, t_{ox}=9 nm (monoclinic HfO_2, EOT=3 nm) and t_{ox}=15 nm (tetragonal HfO_2, EOT=3 nm).
Fig. 11. $I_D(V_D)$ Characteristic of the MOSFET high-$k$ multilayer $\text{HfO}_2/\text{SiO}_2$ (monoclinic-$\text{HfO}_2$ oxide).

Fig. 12. $I_G(V_G)$ Characteristic of the MOSFET high-$k$ multilayer $\text{HfO}_2/\text{SiO}_2$ (monoclinic-$\text{HfO}_2$ oxide).

Fig. 13. $I_D(V_D)$ Characteristic of the MOSFET high-$k$ multilayer $\text{HfO}_2/\text{SiO}_2$ (tetragonal-$\text{HfO}_2$ oxide).

Fig. 14. $I_G(V_G)$ Characteristic of the MOSFET high-$k$ multilayer $\text{HfO}_2/\text{SiO}_2$ (tetragonal-$\text{HfO}_2$ oxide).

The $I(V)$ characteristic of the MOSFET high-$k$ (Fig. 7, Fig. 8) (monolayer $\text{HfO}_2$, monoclinic-$\text{HfO}_2$ oxide) and $I(V)$ characteristic of the MOSFET high-$k$ (Fig. 9, Fig. 10) (monolayer $\text{HfO}_2$, tetragonal-$\text{HfO}_2$ oxide) have the same results but with different thickness ($t_{\text{monoclinic-HfO}_2} = 14$ nm, $t_{\text{tetragonal-HfO}_2} = 22$ nm).

$I(V)$ characteristic of the MOSFET high-$k$ (Fig. 11, Fig. 12) (multilayer $\text{SiO}_2/\text{HfO}_2$, monoclinic-$\text{HfO}_2$ oxide) and $I(V)$ characteristic of the MOSFET high-$k$ (Fig. 13, Fig. 14) (multilayer $\text{SiO}_2/\text{HfO}_2$, tetragonal-$\text{HfO}_2$ oxide) have the same results but with different thickness ($t_{\text{monoclinic-HfO}_2} = 9$ nm, $t_{\text{tetragonal-HfO}_2} = 15$ nm).

From the results obtained (Fig. 7, Fig. 8, Fig. 9, Fig. 10, Fig. 11, Fig. 12, Fig. 13 and Fig. 14), $\text{HfO}_2$ is a good candidate to replace $\text{SiO}_2$.

$I(V)$ characteristic simulated by the analytical model is compared with the neural model. It can be observed that good agreement between the analytical and neural data was obtained by both of the models.

The differences between the results of monoclinic and tetragonal-$\text{HfO}_2$ is:

1. Thickness of tetragonal-$\text{HfO}_2$ is wider than Thickness of monoclinic-$\text{HfO}_2$ (The monoclinic-$\text{HfO}_2$ has a smaller dielectric constant than the tetragonal-$\text{HfO}_2$).

2. The tetragonal-$\text{HfO}_2$ is more stable than the monoclinic-$\text{HfO}_2$.

Table 2 summarizes the percentage error between analytical model and ANN model in different cases.
Table 2
Percentage error between analytical model and ANN model in different cases.

<table>
<thead>
<tr>
<th>Different cases of oxide</th>
<th>Percentage error %</th>
</tr>
</thead>
<tbody>
<tr>
<td>monolayer HfO$_2$, monoclinic-HfO$_2$</td>
<td>1.37 %</td>
</tr>
<tr>
<td>monolayer HfO$_2$, tetragonal-HfO$_2$</td>
<td>1.2 %</td>
</tr>
<tr>
<td>multilayer SiO$_2$/HfO$_2$, monoclinic-HfO$_2$</td>
<td>2.11 %</td>
</tr>
<tr>
<td>multilayer SiO$_2$/HfO$_2$, tetragonal-HfO$_2$</td>
<td>2.24 %</td>
</tr>
</tbody>
</table>

Our objective here is to implement neural network programs in Matlab, using the backpropagation to identify different curves (characteristic I (V) of the MOSFET, the evolution of the average learning error of our predictor). This model can be used to simulate the MOSFET response to different parameters related to the manufacturing technology.

The great advantage of neural networks lies in their automatic learning capability, which allows solving problems without requiring writing complex rules. The results in this work, in good agreement with previous results, [13,14,15,16,17,18]. This last observation shows the applicability of the networks of artificial neurons being studied of nanometric circuits CMOS.

7. Conclusion.
This paper showed the applicability of the neuronal approach to the problem of prediction of the lifespan of the strongly submicron integrated devices. An analytical approach based on a neuronal predictor was developed in the case of transistor MOSFET has high permittivity (HfO). The latter enabled us to envisage the evolution of the current according to the different parameters (drain voltage, gate voltage, channel length and oxide thickness).

8. Reference.

