New Method for Verification of Analytical Device Models Using Transistor Parameter Fluctuations

Christian Kühn¹, Stephan Marksteiner², Thomas E. Kopley²,³ and Werner Weber²

¹Institute of Electronics, ET4, Universität der Bundeswehr, D-85577 Neubiberg, Germany
²Siemens AG, ZT ME, D-81730 Munich, Germany,
³now with Hewlett-Packard Company, Palo Alto, CA

Abstract

In this work device parameter fluctuations are considered for verification of analytical device models. We focus on the \( (WL)^{1/2} \)-dependence of the standard deviation of the threshold voltage \( \sigma_{V_t} \) [1-6]. Besides the known contribution from doping variations, an explicit channel length dependence enhances substantially the \( (WL)^{1/2} \)-curve. Oxide thickness variation is shown to have minor influence on the \( \sigma_{V_t} \)-fluctuation and the contribution of mobility fluctuation is completely negligible for a 0.5 \( \mu \)m process. It has been found that a percolation model [13] for the current paths through the device can give full account of the experimentally measured values of \( \sigma_{V_t} \).

Surface Channel n-MOSFET

In general, for a \( V_T \)-model which is a function of different statistically independent parameters \( p_i \), \( \sigma_{V_t}^2 \) can be expressed as the sum of the squared standard deviations \( \sigma_{V_t,p_i}^2 \) of the corresponding parameters \( p_i \):

\[
V_T = V_T(p_1, \ldots, p_n) \rightarrow \sigma_{V_t}^2 = \sum_{i=1}^{n} \left( \frac{\partial V_T}{\partial p_i} \right)^2 \cdot \sigma_{p_i}^2 , \quad (1)
\]

The following simple \( V_T \)-model is commonly used for the n-MOSFET [7]:

\[
V_T = V_{FB} + \Psi_B + \frac{\sqrt{2e_0\varepsilon_S}}{C_{Ox}} q N_A \Psi_B + \frac{q D_i}{C_{Ox}}, \quad (2)
\]

where \( V_{FB} \) is the flat-band-voltage, \( \Psi_B \) the surface potential at the onset of strong inversion, \( N_A \) the bulk doping, \( D_i \) the threshold adjust implant dose and \( C_{Ox} \) the gate oxide capacitance per unit area.

The analytical calculation of the standard deviation \( \sigma_{V_t} \) yields for the surface channel n-MOSFET [11]:

\[
\sigma_{V_t,\text{doping}} = \frac{1}{C_{Ox}} \sqrt{\frac{q Q_B}{4} + q^2 D_i} \cdot \frac{1}{\sqrt{WL}}, \quad (3)
\]

where \( Q_B \) is the depletion layer charge density and \( WL \) the effective gate area.

Equation (3) suffers from two shortcomings: first, the usual method for determining \( D_i \) is to apply a certain bulk doping level \( N_A \) (determined e.g. from SIMS data) and then to fit \( D_i \) in order to reproduce the measured long-channel value of \( V_T \).

This leads to an ambiguity in the value of the parameter \( D_i \), and thus to an uncertainty in \( \sigma_{V_t} \) according to (3). As an example, we use the doping profile of the n-MOSFET of a 0.5 \( \mu \)m process [5] given in Fig. 1. Using the measured value of \( N_A \) and fitted value of \( D_i \) (to obtain the correct long-channel \( V_T \)) approximately 80 \% of the experimental \( V_T \)

![Figure 1: Doping profile for the NMOS surface channel determined by SIMS. The fitted doping value from the long-channel \( V_T \) is marked by a dotted line.](image-url)
Variations are explained. However, the calculated \( D_i \) is found to be largely exaggerated as compared to the SIMS profile (Fig. 1). On the other hand, optimizing \( N_d \) and \( D_i \) so that the SIMS profile is well represented yields for devices with different \( W \) and \( L \) geometries a \( \sigma_{V_i} \) that is only about 50% of the measured value (Fig. 2).

This result is furthermore confirmed by two-dimensional MEDICI simulations using the exact profile in Fig. 1, which are in good agreement with the more refined 2D simulation method by Stolk and Klaassen [4]. There, the local variation of the doping concentration is considered dividing a cut through the device into cells and using a Poisson distribution for the number of dopants in each cell.

The second shortcoming of (3) is that it cannot account for the channel-length dependence of \( \sigma_{V_i} \) shown in Fig. 3. Here, the short-channel devices show considerably larger \( \sigma_{V_i} \)’s (for the same \( W \times L \)) than the long channel n-MOSFETs. An explanation of this effect can be based on a more sophisticated \( V_T \)-model [9] which considers short channel effects:

\[
V_T = V_{T0} - \eta \left( 2\Phi_B + U_{DS} \right);
\]

\[
\eta = 2\eta_{ox} \exp \left( -\frac{\eta_{ox}}{4x_d} \right),
\]

where \( V_{T0} \) is the threshold voltage of the long channel device, \( \Phi_B \) the built-in-voltage of the p-n-junction between source/drain and the channel, \( T_{ox} \) the oxide thickness, \( x_d \) the depth of the space charge layer and \( L \) the effective channel length. The values of \( \sigma_{V_i} \) resulting from this refined model are also shown in Fig. 3. The channel length effect is clearly visible and the separation of the two branches is in good agreement with the experiment. It can thus be concluded that this channel length effect comes from the reduction of the effective channel length due to the source and drain depletion layers as modeled in (4). However, there is still a difference of approximately 50% between experimental and theoretical data.

Let us now turn to the discussion of additional sources for the variation of \( V_T \) that can account for the remaining 50%.

Variations of the oxide thickness are investigated in [10] by AFM (atomic force microscopy) stating a rms surface roughness \( \Delta = 0.3 \text{nm} \) of the oxide and its correlation length \( L_c = 15 \text{nm} \).

In order to estimate the influence of oxide thickness variations on the \( V_T \)-fluctuation a simplification that is comparable to the one that leads to (3) is used. There, the doping is treated as homogenous for the whole device, neglecting local fluctuations.

Thus an average of the oxide thickness and its standard deviation \( \sigma_d \) is calculated with the use of \( \Delta \) and \( L_c \):

\[
\sigma_d^2 = 2L_c^2\Delta^2 \cdot \frac{1}{WL}.
\]

The factor 2 originates from the assumption that the roughness of both the Si/SiO₂ interface and the SiO₂/gate interface are equal as reported in [11].

Combining the above equation with (1) and (2) leads to an influence of oxide thickness variation on \( \sigma_{V_i} \) of:

\[
\sigma_{V_T,\Delta_{ox}} = \frac{Q_B + qL_c^2}{\varepsilon_0\varepsilon_{Si}} \sqrt{2L_c\Delta} \cdot \frac{1}{\sqrt{WL}}.
\]

Using again the data of our 0.5 \( \mu \text{m} \) process, the contribution of \( \sigma_{V_i,\Delta_{ox}} \) to the total threshold variation \( \sigma_{V_i} \) is found two orders of magnitude smaller than \( \sigma_{V_i,\text{doping}} \) which confirms the observation by Mizuno [2] concerning the independence of \( \sigma_{V_i} \) from oxide thickness variation.

Calculations of \( \sigma_{V_i,\mu} \) due to fluctuations of the mobility are based on the comprehensive mobility model in [12], where the mobility \( \mu(E_T) \) in the inversion layer is determined as a function of the transverse electric field. A requirement for the validity of this model is that the device dimensions are large enough to ensure the absence of channel-length effects.

\[
\mu(E_T) = \mu_0 \left( 1 - \frac{E_T}{E_{ox}} \right)^{\frac{1}{n}}.
\]

This result is further confirmed by two-dimensional MEDICI simulations using the exact profile in Fig. 1, which are in good agreement with the more refined 2D simulation method by Stolk and Klaassen [4]. There, the local variation of the doping concentration is considered dividing a cut through the device into cells and using a Poisson distribution for the number of dopants in each cell.

The second shortcoming of (3) is that it cannot account for the channel-length dependence of \( \sigma_{V_i} \) shown in Fig. 3. Here, the short-channel devices show considerably larger \( \sigma_{V_i} \)’s (for the same \( W \times L \)) than the long channel n-MOSFETs. An explanation of this effect can be based on a more sophisticated \( V_T \)-model [9] which considers short channel effects:

\[
V_T = V_{T0} - \eta \left( 2\Phi_B + U_{DS} \right);
\]

\[
\eta = 2\eta_{ox} \exp \left( -\frac{\eta_{ox}}{4x_d} \right),
\]

where \( V_{T0} \) is the threshold voltage of the long channel device, \( \Phi_B \) the built-in-voltage of the p-n-junction between source/drain and the channel, \( T_{ox} \) the oxide thickness, \( x_d \) the depth of the space charge layer and \( L \) the effective channel length. The values of \( \sigma_{V_i} \) resulting from this refined model are also shown in Fig. 3. The channel length effect is clearly visible and the separation of the two branches is in good agreement with the experiment. It can thus be concluded that this channel length effect comes from the reduction of the effective channel length due to the source and drain depletion layers as modeled in (4). However, there is still a difference of approximately 50% between experimental and theoretical data.

Let us now turn to the discussion of additional sources for the variation of \( V_T \) that can account for the remaining 50%.

Variations of the oxide thickness are investigated in [10] by AFM (atomic force microscopy) stating a rms surface roughness \( \Delta = 0.3 \text{nm} \) of the oxide and its correlation length \( L_c = 15 \text{nm} \).

In order to estimate the influence of oxide thickness variations on the \( V_T \)-fluctuation a simplification that is comparable to the one that leads to (3) is used. There, the doping is treated as homogenous for the whole device, neglecting local fluctuations.

Thus an average of the oxide thickness and its standard deviation \( \sigma_d \) is calculated with the use of \( \Delta \) and \( L_c \):

\[
\sigma_d^2 = 2L_c^2\Delta^2 \cdot \frac{1}{WL}.
\]

The factor 2 originates from the assumption that the roughness of both the Si/SiO₂ interface and the SiO₂/gate interface are equal as reported in [11].

Combining the above equation with (1) and (2) leads to an influence of oxide thickness variation on \( \sigma_{V_i} \) of:

\[
\sigma_{V_T,\Delta_{ox}} = \frac{Q_B + qL_c^2}{\varepsilon_0\varepsilon_{Si}} \sqrt{2L_c\Delta} \cdot \frac{1}{\sqrt{WL}}.
\]

Using again the data of our 0.5 \( \mu \text{m} \) process, the contribution of \( \sigma_{V_i,\Delta_{ox}} \) to the total threshold variation \( \sigma_{V_i} \) is found two orders of magnitude smaller than \( \sigma_{V_i,\text{doping}} \) which confirms the observation by Mizuno [2] concerning the independence of \( \sigma_{V_i} \) from oxide thickness variation.

Calculations of \( \sigma_{V_i,\mu} \) due to fluctuations of the mobility are based on the comprehensive mobility model in [12], where the mobility \( \mu(E_T) \) in the inversion layer is determined as a function of the transverse electric field. A requirement for the validity of this model is that the device dimensions are large enough to ensure the absence of channel-length effects.

\[
\mu(E_T) = \mu_0 \left( 1 - \frac{E_T}{E_{ox}} \right)^{\frac{1}{n}}.
\]
Figure 4: Schematic top view of the oxide - semiconductor interface plane: Due to local inhomogeneities of the doping the surface potential of some parts of the interface has already reached the strong inversion (hatched squares) whereas others are still in the state of weak inversion (blank squares). Depending on the local distribution of the inverted areas percolation paths through the device can appear.

enough to establish an effective carrier mobility from the ensemble averaging of single scattering processes, which is true for the 0.5 µm process. The transverse electric field depends on the doping concentration and is thus a fluctuating parameter. A lengthy calculation shows that for a fixed oxide charge density of $2 \times 10^{10}$ cm$^{-2}$ the contribution from $\sigma V_{t,doping}$ is several orders of magnitude smaller than $\sigma V_{t,el}$. All existing calculations of parameter fluctuations suffer from neglecting the full three-dimensional character of the problem. Percolation occurs due to the spatial variation of the distribution of doping atoms which causes local potential fluctuations at the interface between semiconductor and gate oxide (Fig. 4). This leads to variations in the current paths in the device. An analytic approximation in [13] subdivides the gate area into square elements and assumes those either as conducting or non-conducting. On this basis a formula for $\sigma V_t$ is derived:

$$\sigma_{V_t,perc} = \sqrt{\frac{\pi}{2}} \frac{W}{x_d} \cdot \frac{L}{x_{el}} \cdot \sigma_{Vt,el},$$

where $x_d$ is the length of the square subdivisions of the gate area and $\sigma_{Vt,el}$ is the standard deviation of threshold voltage for a subdivision square. Assuming a normal distribution for the doping in each subdivision square leads to:

$$\sigma_{Vt,perc} = \sqrt{\frac{\pi}{2}} \frac{q}{2 C_{OX}} \left( \frac{N_{A} x_d}{x_{el}} \right)^{1/4} \cdot (WL)^{3/8},$$

where $x_d$ is the depth of the space charge layer below the gate.

For the length of the subdivisions two different values have been used, the depth of the space charge layer, as proposed in [13], and the Debye length, as a typical length for the influence of a potential perturbation in the semiconductor.

Fig. 5 shows the results of these calculations together with the experimental data. As the percolation model has a digital character and thus neglects subthreshold characteristics, it leads to calculated values that are higher than the actually measured ones. We conclude that a percolation model for the current in the subthreshold region can indeed account for the remaining 50% between experimental and theoretical values in Fig. 2.

**Buried Channel p-MOSFET**

Let us now turn to the buried-channel p-MOSFET, where a similar derivation of the $V_T$-variation was carried out:

$$\sigma^2_{V_T} = \left( \frac{\partial V_T}{\partial N_p} \right)^2 \cdot \sigma^2_{N_p} + \left( \frac{\partial V_T}{\partial N_n} \right)^2 \cdot \sigma^2_{N_n},$$

$$\sigma^2_{N_i} = \frac{N_{Ai} + N_{Di}}{x_i} \cdot \frac{1}{WL},$$

where $N_p$ and $N_n$ are the net doping concentrations of the p- and the n-region, $N_{Ai}$ and $N_{Di}$ are the actual acceptor and donor concentrations and $x_i$ is the depth of the space charge layer in the respective region.

This general formula is applied to the different $V_T$-models proposed by Sze [7], Van der Tol [14] and Klaassen [15]. The model of Van der Tol can be converted into the one suggested by Nishiuchi [16] and adopted by Sze.

For our calculations on the p-MOSFET, values for the average doping concentrations are again taken from SIMS data (Fig. 6). The calculated standard deviations of the threshold voltage according to (9) and (10) using Sze's formula cannot be brought into agreement with the 50% disparity between experiment and corresponding simple theory (3) for the NMOS, that assumes an overall distribution for the quantity of dopant atoms. The contributions to the standard deviation from additional effects as discussed above
Applying (1) to his $V_T$-formula we get values for the standard deviation from the direct influence of doping fluctuations without consideration of percolation effects and channel length effects, that reach about 50% of the experimental data similar to the NMOS case. In Fig. 7 the results of these calculations are shown together with the theoretical curves for the percolation model applied to the PMOS case.

In conclusion the PMOS model of Klaassen leads to good agreement between our theoretical considerations and the experimental data and is thus more favorable than that of Sze. This demonstrates, that device models can be tested based on the measurement of device parameter variations.

**Acknowledgment**

The authors wish to thank K. Hoffmann (Universität der Bundeswehr) for his contribution to this work. This work was supported by the DFG (Ho 1325/2-2).

**References**


