Local oxide capacitance as a crucial parameter for characterization of hot-carrier degradation in long-channel n-MOSFETs

Ivan Starkov
Institute for Microelectronics, Vienna University of Technology, Gußhausstraße 27-29, A-1040 Vienna, Austria

Hubert Enichlmair
Process Development and Implementation Department, Austriamicrosystems AG, Tobelbader Straße 30, A-8141 Unterpremstätten, Austria

(Received 13 August 2012; accepted 19 December 2012; published 9 January 2013)

A theoretical model for the MOSFET local oxide capacitance as a crucial parameter for the characterization of hot-carrier degradation has been developed. For this purpose, the conformal mapping technique is used. On the basis of the proposed approach a refined extraction scheme for the defect distribution from charge-pumping measurements has been employed. Assuming the extracted spatial trap distributions at different stress times as input, the transfer characteristics and linear drain current degradation are numerically calculated and compared with the experimental results. A very good agreement is achieved. These results demonstrate that the coordinate dependence of the oxide capacitance is extremely important for an accurate extraction of the defect profile particularly for large stress times. Additionally, the obtained results confirm the findings of our physics-based model of hot-carrier degradation. © 2013 American Vacuum Society.

I. INTRODUCTION

Hot-carrier degradation (HCD) is associated with the build-up of interface states and oxide trapped charges (with densities $N_{it}$ and $N_{oa}$) of an metal-oxide-semiconductor field-effect-transistor (MOSFET). Therefore, quantitative information on the defect spatial distributions is essential to reveal and understand the physical mechanisms of the HCD phenomenon. For this purpose, the charge pumping (CP) technique is widely used.1–4 Most methods for extraction of the lateral defect profiles from CP data employ a constant transistor oxide capacitance, namely

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox,0}},$$

(1)

where $t_{ox,0}$ is the oxide thickness at the center of the device and $\varepsilon_{ox}$ is the dielectric permittivity (e.g., Refs. 5 and 6). This approach is applicable only for transistors with a fixed oxide thickness and leads to erroneous results when $t_{ox}$ depends on the interfacial coordinate $x$. Moreover, even in the case of a uniform thickness, such a scheme leads to spurious evaluation due to the neglect of fringing fields. Based on the work of Chim et al.,4 we propose an extension of Eq. (1) when $t_{ox}$ depends on $x$ ($t_{ox} = t_{ox}(x)$). Even in such an approach, the MOS structure is considered as an ideal infinite parallel-plate capacitor or, in other words, the oxide electric field is assumed uniform. In practice, however, a substantial distortion occurs near the source/drain ends of the gate. The aforementioned simplification does not strongly affect the transistor characteristics. Nevertheless, accounting of the fringing effect is of great importance for the characterization of the defect spatial distributions after hot-carrier stress because the $N_{it}(x)$ peak is located near the drain end of the gate1–5 where the electric field nonuniformity is most pronounced. Among the great number of publications devoted to $N_{it}(x)$ and $N_{oa}(x)$ extraction algorithms, we were able to find consideration of the oxide capacitance coordinate dependence only in the work of Lee et al.3 This approach is based on the simulation of the local threshold or flatband voltage shift induced by the given uniform oxide charge distribution and requires adequate computational resources. Thus, a compact analytical model for simplification of the defect profile extraction technique is of great importance even nowadays. For the solution of this problem, we use the conformal-mapping method, which is most helpful for analysis of the fringing electric field effect under simple two-dimensional boundary conditions.2

II. EXPERIMENTAL SUPPORT

To validate the developed analytical approach, we used two 5 V n-MOSFETs with identical architecture (Fig. 1) differing only in channel lengths ($L_{ch} = 0.5$ and 2.0 $\mu$m). The drain-sided gate edge is the origin of the $x$-axis for both devices. Transistors were fabricated on a standard 0.35 $\mu$m technology and subjected to a hot-carrier stress at the gate voltage of $V_{gs} = \{2.0, 2.6\}$ V and the drain voltage of $V_{ds} = \{6.25, 6.5\}$ V up to 10$^4$ s. The oxide thickness of the transistor, which can be obtained from the process simulator, is interpreted as $t_{ox,1}(x) = t_{ox,0} + \Delta t_{ox}(x)$, where $\Delta t_{ox}(x)$ is the thickness gradient component, which increases closer to the end of the gate contact (Fig. 2). Additionally, to ensure that the device geometrical peculiarities are taken into account, the proposed model was examined on two artificial devices with $L_{ch} = 0.5 \mu$m but with different oxide thicknesses of $t_{ox,0}(x) = t_{ox,0}$ and $t_{ox,2}(x) = t_{ox,0} + 2\Delta t_{ox}(x)$, see Fig. 2.

For charge-pumping current measurements, we use an experimental scheme proposed in Ref. 4. The gate of the transistor is connected to a pulse generator, while source, drain,
and substrate contacts are grounded. As suggested in Refs. 2 and 4, we measured the CP current at the drain, source, and substrate. Due to the presence of the damage dose provided by the multiple-carrier component of the Si-H bond dissociation process, it is important to perform separate current measurements. The gate of the MOSFET was pulsed by a trapezoidal waveform at a frequency of $f = 100$ kHz. We use $V_{gs} = 5$ V and increase $V_{gs}$ from $-4$ to $3$ V (varying high-level technique) and $V_{gs} = 4$ V with decreasing $V_{gs}$ from $3$ to $-5$ V (varying low-level technique), where $V_{gs}$ and $V_{gh}$ are the low and the high levels of the gate pulse, respectively. In order to obtain sufficient spatial resolution of the defect profiles, the voltage step was set to 0.02 V.

III. CONFORMAL-MAPPING METHOD

The conformal-mapping method is most helpful for consideration of fringing electric fields in simple two-dimensional boundary problems (which is just our case) by transforming the boundary to a soluble form. A series of simulations based on the Lee et al. approach (described below) allow us to come to a conclusion that for a local oxide capacitance consideration a simplified structure can be used. Namely, the gate contact can be interpreted geometrically as a ray instead of more complicated corner variant. The problem with the coordinate system is shown in Fig. 3. Here, the origin of the considered system is placed at the drain end of the gate. Additionally, to be able to apply the analytical technique, it is necessary to introduce several assumptions: (i) the oxide can be considered as homogeneous and isotropic media; (ii) the silicon interface is equipotential at the potential $U = U_0$, in other words, there is no potential discontinuity from the highly doped drain contact region to the middle of the device channel; (iii) the gate interface is equipotential with the zero potential $U = 0$; and (iv) the oxide thickness is much smaller than the space between the gate and drain contacts.

The $z = x + iy$ plane is mapped into the $\omega = \varphi + i\psi$ plane with the functional relationship between $z$ and $\omega$ described by

$$z = \frac{t_{ox}}{\pi} \left[ \omega + \exp(\omega) \right],$$

or in the coordinate notations

$$x = \frac{t_{ox}}{\pi} \left[ \varphi + \exp(\varphi) \cos \psi \right],$$

$$y = \frac{t_{ox}}{\pi} \left[ \psi + \exp(\varphi) \sin \psi \right].$$

The suggested conformal transformation in Fig. 3 reduces the original problem to the Laplace problem between two parallel infinitely long metallic plates at different potentials. The analytical solution for the potential and the electric field distribution in the oxide near the gate end following from the suggested conformal map and that calculated by the device and circuit simulator MiniMOS-NT are presented in Fig. 4. One can see that the simulation results confirm the analytical calculations with a rather good accuracy, which means correctness of chosen geometrical system interpretation.

The local oxide capacitance can be defined as the ratio between the surface charge density $\sigma_{surface}$, and the interface potential $U_0$ and can be written as

Fig. 1. (Color online) Topology of the n-MOSFET ($L_{ch} = 0.5\mu m$) with the net doping profile highlighted.

Fig. 2. (Color online) Dependence of the device oxide thicknesses vs the lateral coordinate. Additionally, the position of the carrier acceleration integral maxima calculated for real and artificial devices are shown.

Fig. 3. (Color online) Conformal transformation used to solve the gate/drain fringing problem. The complicated case of the corner gate form (red dashed line) is reduced just to a ray.
It should be mentioned that in Eq. (10), we have changed the parameter \( t_{ox} \) to \( t_{ox}(x) \). As reported in Ref. 11, this substitution is legitimate until \( d_{ox}(x)/dx < 1 \), which is correct for devices with a real topology of the gate contact. The value of \( C_{ox}(x) \) in Eq. (10) has a maximum with \( C_{ox}(= \varepsilon_{ox}/t_{ox}) \) in the middle of the gate and decreases gradually toward a much lower value outside the gate edge due to the fringing effect. At \( x \to -\infty \) (or \( \phi \to -\infty \)), the obtained Eq. (10) asymptotically turns into well known expression for the parallel-plate capacitance (1).

IV. APPLICATION OF THE DEVELOPED APPROACH

It is rather important to clarify at which distance from the drain edge of the gate contact the electric field becomes uniform. In other words, the location of where the fringing effect is negligible and the parallel-plate capacitor approximation is applicable. For this purpose, let us find a position at the device interface (\( x \)-axis) where the electric field intensity \( F \) differs by 1% from the uniform one, i.e., \( F_0 = U_0/t_{ox} \) (e.g., Ref. 7). For any position of the structure considered in Fig. 3, the electric field intensity can be defined using the conformal transformation\(^7\) as \( F = |d\omega/dz| \). Therefore, from Eq. (8) one may conclude

\[
F = U_0 \frac{1}{t_{ox} \cdot 1 + \exp(\psi)},
\]

or in terms of \( F_0 \)

\[
\frac{F}{F_0} = \frac{1}{1 + \exp(\varphi)}.
\]

Here, the same transition as from Eq. (7) to Eq. (8) is used. Because, we are interested in the electric field distribution at the silicon interface (or at the lower infinite plate) one assumes \( \psi = 0 \). Assuming \( F/F_0 = 0.99 \), we obtain \( \exp(\varphi) = 0.0101 \) and, as a consequence, \( \varphi = -4.61 \). Substitution of the obtained values in the expression for the \( x \) coordinate in Eq. (3) results in

\[
x_0 = 4.61 \frac{t_{ox}(x)}{\pi} = 1.47 t_{ox}(x = L_g).
\]

Here, \( x = L_g \) is the position of the drain end of the gate (in our case \( L_g = 0 \)). Thus, one can consider the local oxide capacitance as constant, i.e., use the parallel-plate capacitor approximation, already at a distance of 1.47 of the oxide thickness at the drain end of the gate contact. In other words, it is possible to operate with a simple expression (1) if the region of interest is outside of the mentioned area.

V. LEE’S APPROACH FOR MODELING THE LOCAL OXIDE CAPACITANCE

For the evaluation of the oxide capacitance coordinate dependence using device modeling, we employ the method developed by Lee et al.\(^3\) To determine \( C_{ox}(x) \) by means of simulation, a careful calculation of the local flatband \( V_{fb}(x) \)
and threshold $V_{th}(x)$ voltage distributions is performed. Due to the symmetry of the source and drain for the fresh device, we present results only for the drain half of the device. As mentioned in Ref. 1, for an unstressed transistor, $V_{th}(x)$ and $V_{fb}(x)$ profiles can be obtained employing device simulator MiniMOS-NT. The carrier distribution required for the local threshold and flatband voltage calculation was estimated to be $5.1 \times 10^{13}$ cm$^{-3}$ and $1.6 \times 10^{14}$ cm$^{-3}$ for the simulation$^{1,12}$ respectively. The soundness of the obtained result has been checked by Monte-Carlo simulation.$^{13}$ Figure 5 demonstrates that the Monte-Carlo curves (symbols) practically coincide with those obtained from MiniMOS-NT (solid lines). Note that in the low voltage region the agreement is worse due to the stochastic nature of the Monte-Carlo approach, which cannot provide sufficient accuracy in this area. For any point at the device interface, the poststress local threshold $V_{th, s}(x)$ and local flatband $V_{fb, s}(x)$ values are related to the prestressed ones by$^{1,5}$

$$V_{th, s}(x) = V_{th}(x) - \frac{q\Delta N_{ox}(x)}{C_{ox}(x)} + \frac{q\Delta N_{fb}(x)}{2C_{ox}(x)}$$

$$V_{fb, s}(x) = V_{fb}(x) - \frac{q\Delta N_{ox}(x)}{C_{ox}(x)} - \frac{q\Delta N_{fb}(x)}{2C_{ox}(x)},$$

(14)

which considers the change between pre- and poststress concentrations of interface traps and bulk oxide charges. The presence of a probe uniform oxide charge $N_{ox, uniform}$ leads to a local threshold voltage shift $\Delta V_{th, uniform}(x) = V_{th, s(uniform)}(x) - V_{th}(x).$ Therefore, the local oxide capacitance can be found as$^{3}$

$$C_{ox}(x) = -\frac{N_{ox, uniform}}{\Delta V_{th, uniform}(x)}.$$  

(15)

A typical example of $\Delta V_{th, uniform}(x)$ induced by a uniform oxide charge density of $5 \times 10^{11}$ cm$^{-2}$ is shown in Fig. 5.

VI. RESULTS AND DISCUSSION

A. Long-channel devices

A comparison of the simulation approach$^3$ and the newly developed analytical model (10) for the local capacitance profile $t_{ox}(0.1, 2)$ is presented in Fig. 6. One may conclude that the obtained results are in good agreement. At the same time, the expression $C_{ox} = \varepsilon_{ox}/t_{ox}(0.1, 2)$ leads to substantially different results. The value of $C_{ox}(0.1, 2)$ has its maximum with $C_{ox,0}(= \varepsilon_{ox}/t_{ox,0})$ in the middle of the gate and decreases gradually toward a much lower value outside the gate edge due to the fringing effect. As expected, the most pronounced peculiarity is observed at the drain side of the gate where the abrupt change in the oxide thickness occurs, Fig. 2. From a detailed analysis of suggested conformal transformation, we conclude that under the gate electrode fringing effect can be neglected for distances larger than $1.47 t_{ox}(x = L_2)$ from the gate edge (this is confirmed in Fig. 6 for all $t_{ox}(0.1, 2)$). Note that an abrupt reduction in $C_{ox}(x) = \varepsilon_{ox}/t_{ox}(0.1, 2)$ at $x = 0 \mu$m is unphysical and such an approach should not be used.

The derived analytical expression for the local oxide capacitance (10) was incorporated into the Chim et al. extraction scheme. With this refined characterization approach, we were able to extract the evolution of the $N_{ox}$ densities with the stress time from CP measurements and compare with the results calculated using the parallel plate capacitor approximation (1). The defect profiles obtained under different assumptions on the $C_{ox}(x)$ distribution for $V_{th} = 6.5$ V and $V_{gs} = 2.6$ V at the stress time of $10^5$ s are presented in Fig. 7. For this figure, we incorporated $C_{ox}(x) = \varepsilon_{ox}/t_{ox,0}$ and the proposed model (10) for $t_{ox, 1}$. The extracted $N_{ox}(x)$ and $N_{th}(x)$ concentrations were subjected to further validation as input parameters to simulate the transfer characteristics of the degraded device employing MiniMOS-NT. Comparison of simulated and experimental curves once again confirms the applicability of the developed model for...
the local oxide capacitance (see Fig. 8). The role of the fringing effect at large stress times is obvious.

Additionally, the evolution of defect profiles with stress time was extracted from CP data for the stress conditions of \( V_{ds} = 6.25 \text{ V} \) and \( V_{gs} = 2.0 \text{ V} \). Extraction procedure results calculated with \( C_{ox,1}(x) \) are demonstrated in Fig. 9 for both devices under consideration. One can see in Fig. 9 that the obtained defect profiles feature two peaks starting from \( \sim 10^5 \text{s} \). Moreover, Fig. 9 (in Fig. 7 depicts the same tendency) demonstrates that these peaks just correspond to the maxima of the electron and hole acceleration integrals. This result is confirmed by the findings of our HCD model, which shows that these peaks are related to the contributions induced by primary channel electrons and secondary holes generated by impact ionization caused by the injection of hot electrons.

Our modeling approach is based on the information how efficiently carriers trigger the bond dissociation process. We consider the superposition of single- and multiple-carrier mechanisms of Si-H bond-breakage for both channel electrons and holes. These mechanisms are controlled by the carrier acceleration integral (AI)

\[
AI_{(e,h)} = \int_{E_{th}}^{\infty} f_{(e,h)}(E)g(E)\sigma_{(e,h)}(E)v(E)\mathrm{d}E,
\]

which is calculated as the reaction cross section \( \sigma_{(e,h)} \) (are the Keldysh-like reaction cross sections) multiplied on the density-of-states \( g(E) \) and the carrier velocity \( v(E) \), weighted with the carrier energy distribution function \( f(E) \) and integrated over energy starting from the threshold value typical for the bond dissociation reaction. As we demonstrated in Ref. 8, the nonuniform nature of HCD in long-channel devices is related to the single-carrier bond-breakage process. In literature, one may find different criteria how efficiently carriers interact with the Si-H bond, e.g., the maximum of the electric field, the carrier dynamic temperature, position where the distribution function demonstrates most extended high-energy tails, etc. However, as we showed in Refs. 8, 14, and 15, the \( N_{it} \) peak just corresponds to the maximum of the AI and is shifted with respect to maxima of other quantities. The acceleration integral has the same functional form for single- and multiple-carrier mechanisms and for both types of carriers. Holes are accelerated by the electric field towards the source, thereby creating interface states shifted with respect to the electron-induced ones. As a result, the \( N_{it} \) fraction induced by holes should be much less than their relative contribution to the linear drain current change because the hole contribution is considerably shifted toward the source. The contribution of channel holes to the total defect concentration is much less than the electron corresponding fraction. This trend becomes more pronounced for longer devices (Fig. 9), see also Refs. 14 and 15.
Trap distributions demonstrated in Fig. 9 were used as input parameters to numerically calculate the linear drain current $I_{dlin}$ degradation employing MiniMOS-NT. After comparison of simulated and experimental data in Fig. 10 once again one may conclude that the proposed model for the local oxide capacitance is important for a proper defect profile extraction. It should be noted that the change of the $I_{dlin}$ degradation slope appearing at $\sim 10^5$ s for both devices can be linked to the contribution of the hole $N_H$ peak to the total defect density.

Both $I_{dlin}$ degradation curves were reproduced by our HCD model. Moreover, we applied our model for devices with artificial oxide thickness, namely for $t_{ox,0}$ and $t_{ox,2}$. Figure 10 depicts that for the device with the flattened gate interface the degradation rate is lower in comparison with the nominal one. Such a tendency can be explained by the difference in the form of high-energy tails of the carrier distribution function and, as a consequence, by the shift of the electron and hole AIs, as shown in Fig. 2. The maximum value of carrier acceleration integrals for the artificial device with $t_{ox,0}$ are shifted in respect to the nominal device, i.e., with real oxide thickness. To be more precise, in the case of electrons, it is around 100 nm.

Note, due to the fringing effect the proposed model is more accurate even without information on the $t_{ox}(x)$ distribution (i.e., only the effective oxide thickness is known from the manufacturing process). For this case, a constant oxide thickness assumption (i.e., $C_{ox,0}$) results in an error $I_{dlin}$ degradation (by incorporating the defect spatial distributions into the device simulator) for stress times longer than 10 ks around 3%–5% while using (1) gives 15%–20%.

B. Scaled devices

In contrast to long-channel devices, for scaled devices with small operating voltages and corresponding topologies the application of the proposed model most probably is not necessary because the $N_0(x)$ peak is located closer to the device center. Confirmation of this fact can be found in Fig. 11 where the evolution of the $N_H$ profile with stress time is demonstrated for an n-MOSFET with $L_{ch} = 0.18 \mu m$. The effective channel length$^4$ ends under the edge of the application area. That is, the standard formula for the parallel-plate capacitor (1) can be used.

VII. CONCLUSION

We have shown that the accurate consideration of the oxide capacitance dependence on the lateral coordinate is essential for a proper extraction of the defect profiles from charge pumping data for long-channel devices. The presented analytical model for $C_{ox}(x)$ was verified by representing transfer characteristics and $I_{dlin}$ degradation in 5 V n-MOSFETs with various channel lengths. We have demonstrated that by ignoring the spatial variation of $C_{ox}(x)$, a spurious result is produced, leading to an ambiguous picture of HCD with the model ignoring the $C_{ox}(x)$ distribution. The obtained results are supported by our model of hot-carrier degradation, which considers a contribution of channel electrons as well as secondary holes generated by impact ionization. Our study may provide a theoretical basis and physical insights for the further refinement of conventional extraction techniques. Moreover, our analysis could also be useful for other problems that consider fringing effect, e.g., capacitance modeling.

---


12. A. N. Tikhonov and A. A. Samarskii, Equations of Mathematical Physics (Dover, New York, 2011).


