A New Approach to Simulating n-MOSFET Gate Current Degradation by Including Hot-Electron Induced Oxide Damage

Cherng-Ming Yih, Shui-Ming Cheng, and Steve S. Chung, Senior Member, IEEE

Abstract—A new gate current model which considers the hot-electron induced oxide damage in n-MOSFET’s was developed for the first time. The spatial distributions of oxide damage, including the interface state \(N_{it}\) and oxide trapped charge \(Q_{ox}\), were characterized by using an improved gated-diode current measurement technique. A numerical model feasible for accurately simulating gate current degradation due to the stress generated \(N_{it}\) and \(Q_{ox}\) has thus been proposed. Furthermore, the individual contributions of \(N_{it}\) and \(Q_{ox}\) to the degradation of gate current can thus be calculated separately using these oxide damage. For devices stressed under maximum gate current biases, it was found that the interface state will degrade the gate current more seriously than that of the oxide trapped charge. In other words, the interface states will dominate the gate current degradation under \(I_{G, max}\). Good agreement of the simulated gate current has been achieved by comparing with the measured data for pre-stressed and post-stressed devices. Finally, the proposed degradation model is not only useful for predicting the gate current after the hot-electron stress, but also provides a monitor that is superior to substrate current for submicron device reliability applications, in particular for EPROM and Flash EEPROM devices.

I. INTRODUCTION

CHANNEnEL-HOT-ELECTRON INJECTION (CHEI) into the gate oxide has been recognized as a major limitation to the long term reliability of short channel n-MOSFET devices. In particular, it has also been widely used as the programming method in EPROM’s and Flash EEPROM’s. Therefore, an understanding of the hot electron related issues becomes of great importance to the device design and reliability study. In recent years, many experimental techniques and theoretical models have been developed to explore and investigate the physics of CHEI through gate leakage current measurement [1]–[3]. However, most of them are focused on the modeling of devices without incurring any hot-electron stress. In other words, the effects of hot-electron stress-induced oxide damage, including interface states \(N_{it}\) and oxide trapped charges \(Q_{ox}\), on the gate current have not been reported in the literature. The difficulties for developing a gate current model of MOS devices after the stress are twofold. One is the difficulty in determining stress generated oxide damage. The other one is the unknown correlation between the gate current and generated oxide damage.

In terms of the oxide damage characterization, many efforts have been devoted to determine their spatial distributions [4]–[6]. But, most of them are focused on the characterization of the localized \(N_{it}\) near the drain region only [4]. Quite a few studies [5]–[6] are able to simultaneously extract \(N_{it}\) and \(Q_{ox}\). Recently, we have successfully demonstrated a method to separate \(Q_{ox}\) from \(N_{it}\) such that the study of these two oxide damage on the gate current degradation becomes feasible [7]. As a consequence, the investigation of the gate current degradation due to hot-carrier stress becomes possible.

On the other hand, to investigate the degradation mechanism of gate current after the stress, Marchetaux et al. [8] observed that the gate current decreases with the aging time. However, due to a lack of the damage profiling technique, the degradation mechanism and model of gate current have never been presented. More recently, the degradation mechanism of gate current and its impact on Flash EEPROM devices, such as the cell programming speed retardation and the operation window closure, have been proposed by an indirect experimental observation [9]. Furthermore, Peng et al. [10] proposed a physically-based interface state and oxide charge generation models for simulating the gate current after the stress. However, both of them cannot explain the experimental data in [8].

In this paper, a new gate current degradation model which considers the hot-electron stress generated \(N_{it}\) and \(Q_{ox}\) will be proposed. The spatial distributions of both \(N_{it}\) and \(Q_{ox}\) will be first calculated using an improved gated-diode current measurement technique that we developed in [7]. Then, the individual contributions of \(N_{it}\) and \(Q_{ox}\) to the degradation of gate current will be studied. Section II describes the device conditions used in this work. The technique for characterizing the oxide damage is described in Section III. The proposed new gate current model is shown in Section IV. Results and discussion are presented in Section V. Section VI gives the summary and conclusions.

II. DEVICE PREPARATION

The devices used in this study were fabricated by 0.35-\(\mu\)m CMOS technology. The test samples have mask length...
III. OXIDE DAMAGE PROFILING TECHNIQUE

In order to investigate the gate current characteristics after the hot-carrier stress, the stress-generated \( N_k \) and \( Q_{\text{ox}} \) should be determined first. We used an improved gated-diode measurement technique that we developed recently in [7] to separate \( N_k \) and \( Q_{\text{ox}} \). Its basic principle will be summarized as follows.

We first stressed the device at the maximum gate current bias since the gate current at this bias is of great interest to us for practical applications such as nonvolatile memories which use this maximum gate current for programming purposes. Under such bias stress condition, both \( N_k \) and \( Q_{\text{ox}} \) will be generated [7]. The measurement method is based on the gated-diode current measurement technique as illustrated in Fig. 1(a). By applying a small negative voltage on the drain side with source floating, and by sweeping the gate voltages from zero to more negative values, the gated-diode current can be measured from the drain. The notations \( \phi_e \) and \( \phi_h \) shown in Fig. 1(a) represent the quasi-Fermi levels for electrons and holes respectively, which coincide with the intrinsic level \( E_i \). The region \( \Delta x \) at the surface between \( \phi_e \) and \( \phi_h \) shows where electron and hole recombination occurs. For a certain applied gate voltage, the region \( \Delta x \) is dependent on the drain voltage. The more negative the applied drain voltage, the larger the electron/hole recombination region becomes. Therefore, the measured \( I_{\text{GD}} \) is easier to be observed by using more negative drain voltage without sacrificing the accuracy. But, Speckbacher et al. [11] suggest that the applied drain voltage greater than \(-0.3 \) V is better. In this paper, the applied drain voltage of \(-0.2 \) V is used. For a fresh device, its \( I_{\text{GD}} \) is shown in Fig. 1(b) as solid lines. After the maximum gate current stress, \( I_{\text{GD,DEG}} \) both \( N_k \) and \( Q_{\text{ox}} \) will be generated, and an excess recombinaton of electron and hole in the region \( \Delta x \) will give rise to an increasing \( I_{\text{GD}} \) (as shown in dashed lines of Fig. 1(b)). In the meantime, the generated \( Q_{\text{ox}} \) will not contribute to the recombination current but will affect the parameters \( \phi_e \) and \( \phi_h \) at the surface. In order to separate \( Q_{\text{ox}} \) from \( N_k \), we use a detrapping or neutralization technique to achieve this purpose, i.e., by applying hot hole injection from the drain to the gate oxide, until \( Q_{\text{ox}} \) was filled and eliminated totally. The \( I_{\text{GD}} \) current will shift to the left [as shown in solid circles in Fig. 1(b)] with a negative oxide trapped charge in this study. If \( Q_{\text{ox}} \) is totally eliminated, the \( I_{\text{GD}} \) current difference between fresh (solid lines) and detrapped (solid circles) curves is caused by \( N_k \) alone. The difference between the stressed (dashed lines) and detrapped/neutralized (solid circles) \( I_{\text{GD}} \) currents can be used to determine the values of \( Q_{\text{ox}} \).

IV. GATE CURRENT DEGRADED MODEL

The measured gate and substrate currents as a function of stress time are shown in Fig. 2. The stress conditions are biased at the maximum gate current (\( I_{\text{G,max}} \) at \( V_{\text{GS}} = 5.5 \) V and \( V_{\text{DS}} = 5 \) V) and at the maximum substrate current (\( I_{B,max} \) at \( V_{\text{GS}} = 2 \) V and \( V_{\text{DS}} = 5 \) V), respectively. The gate current (\( I_G \)) is measured at \( V_{\text{GS}} = 5.5 \) V and \( V_{\text{DS}} = 5 \) V.
The substrate currents \(I_B\) are measured at \(V_{DS} = 2\) V and \(V_{DS} = 5\) V. As seen from Fig. 2, we observe that \(I_B\) increases slightly with the stress time but \(I_C\) decreases. The increase of \(I_B\) is attributed to the enhancement of local electric field as well as the electron energy caused by the generation of \(N_e\) and \(Q_{ox}\). However, it cannot explain the decay of \(I_C\) [3], [10]. Also, we found that the variation of both \(I_G\) and \(I_B\) after \(I_{B,max}\) stress is significantly greater than that after \(I_{G,max}\) stress. Here, it should be noted that at \(I_{B,max}\) stress condition \(Q_{ox}\) will not be generated. Therefore, the generation of \(N_e\) at \(I_{B,max}\) as well as \(I_{G,max}\) will cause an apparent degradation of \(I_G\). Reported results [10] did not provide the correlation between gate current degradation and \(N_e/Q_{ox}\). Therefore, here we propose a new degradation model of gate current by considering the effects of both \(N_e\) and \(Q_{ox}\) as follows.

The generation of channel-hot-electron injection induced gate current includes several steps: 1) channel electrons gain energy from field and become “hot”; 2) redirection of hot electrons due to the scattering; 3) hot electrons surmount the Si/SiO\(_2\) potential barrier and are injected into the oxide; and finally 4) hot electrons are collected by the gate. According to this, the gate leakage current model for devices without considering any hot-electron stress can be described by [3]

\[
J_G = \int_{q_n}^{\infty} qn\nu_z(u_n)f(u_n)au_n
\]  

where \(q\) is the electronic charge, \(n\) is the electron concentration, and \(\nu_z(u_n)\) and \(f(u_n)\) are the electron velocity in the direction normal to the Si/SiO\(_2\) interface and the electron energy distribution function, respectively. As \(N_e\) and \(Q_{ox}\) are generated after the stress, they will increase the lateral electric field as well as the electron energy (from the increase of substrate current). Therefore, \(\nu_z(u_n)\) and \(f(u_n)\) after the stress will be larger than those before the stress. Carrier concentration will also be reduced as a result of \(N_e\) which can be supported by 2-D simulation. As a result, for the increase of \(u_n\) and the reduction of carrier concentration due to \(N_e\), the simulated gate current will increase as described in [10, eq. (3)]. However, it is in contradiction to the measured data in Fig. 2 here. And hence, it is believed that there is another factor responsible for the degradation of gate current as seen from Fig. 2.

In Fig. 3, the measured gate currents after \(I_{B,max}\) stress have shown as a function of charge-pumping current variation (or interface states \(N_e\)). According to Fig. 3 and the correlation of \(I_C\) and \(N_e\), \(\Delta I_C = q\Delta N_e\), the correlation between gate current and \(N_e\) is drawn as an exponential relationship. The factor \(e^{-\alpha_{st}\Delta N_e}\) is proposed and regarded as Coulomb scattering due to trapped interface-state charges as described in [12]. In other words, the generated interface states filled with the electrons can be served as Coulomb scattering centers and then suppress the hot-electron injection capability. In addition, the generated \(Q_{ox}\) trapped in the oxide after \(I_{G,max}\) stress will also inhibit the hot-electron injection probability. Here, we consider the effects of generated \(Q_{ox}\) on the changes of the effective potential barrier height of hot-electron injection at the Si/SiO\(_2\) interface. Therefore, the gate current model which considers the effects of generated \(N_e\) and \(Q_{ox}\) can be formulated as

\[
J_G = \int_{q_n}^{\infty} qn\nu_z(u_n)f(u_n)\exp(-\alpha_{st}\Delta N_e(x))du_n
\]  

where \(\alpha_{st}\) is a fitting parameter for the gate current simulation by considering the \(N_e\) effect. The effective potential barrier height \(\Phi_b\), which considers the stress generated \(Q_{ox}\) can be modified as

\[
\Phi_b = 3.2 - 2.59 \times 10^{-4}E_{ox}^{1/2} - 4.0 \times 10^{-5}E_{ox}^{1/3} + \frac{qQ_{ox0}}{E_{ox}}
\]

The quantity 3.2 V is the Si/SiO\(_2\) interface barrier height. The second term in (5) represents the barrier lowering effect due to the image field. The third term in (5) accounts phenomenologically for the finite probability of tunneling between the silicon and the silicon dioxide. The spatial distributions of both stress-induced \(N_e(x)\) and \(Q_{ox}(x)\) for gate current simulation can be determined by the modified gated-diode current measurement technique [7].
V. RESULTS AND DISCUSSION

A. Characterization of \( N_{It} \) and \( Q_{ox} \)

Fig. 4(a) and (b) show the measured gated-diode currents \((I_{GD})\) and gate-induced-drain-leakage currents \((I_{GIDL})\) for fresh, hot-electron stressed, and detrapped/neutralized devices. Here, \( I_{GD} \) currents will be used to calculate \( N_{It} \) and \( Q_{ox} \), while \( I_{GIDL} \) currents will be used as a monitor for the generation and elimination of \( Q_{ox} \) after the hot-electron stress and after the detrapping/neutralization step. The hot-electron stress, \( I_{G,max} \), stress, is performed at the bias condition of \( V_{GS} = 5.5 \text{ V} \) and \( V_{DS} = 5 \text{ V} \) for 1000 s. Under this bias condition, both \( N_{It} \) and negative \( Q_{ox} \) will be generated. The detrapping/neutralization step for eliminating \( Q_{ox} \) is performed at \( V_{D} = 3 \text{ V} \) and \( V_{G} = -4.5 \text{ V} \) for 50 sec. After this detrapping process, the \( I_{G,max} \) of a stressed device is moved from dashed curve to the right (in solid circles) until the current matches with that of fresh devices at high \( V_{G} \) (e.g., \( V_{G} > 1.2 \text{ V} \)). This means that the stress generated \( Q_{ox} \) is totally detrapped/neutralized. Therefore, the difference of \( I_{GD} - V_{G} \) curves between the fresh (curve 1) and detrapped (curve 3) devices is caused by \( N_{It} \) alone and can be used to determine \( N_{It} \) by (1). Furthermore, the differences of \( I_{GD} - V_{G} \) curves between the detrapped curve (curve 3) and hot-electron stressed curve (curve 2) can be used to calculate \( Q_{ox} \) by (2). Fig. 5 shows the extracted \( N_{It} \) and \( Q_{ox} \) for different stress times. We see that \( N_{It} \) and \( Q_{ox} \) are both localized in the gate–drain overlap region. In addition, the maximum value of stress-induced \( Q_{ox} \) is about four times greater than that of stress-induced \( N_{It} \) at the maximum gate current stress bias.

B. Simulation of Gate Current and the Comparison with Experiment

By incorporating the extracted lateral distributions of both \( N_{It} \) and \( Q_{ox} \) as given in Fig. 5 into a 2-D simulator—ATLAS2 [13], both the device characteristics before and after the stress can be simulated. Fig. 6 shows the simulated lateral electric field distributions at the Si/SiO\(_2\) interface for fresh and after 1000-s stressed devices at \( V_{GS} = 5.5 \text{ V} \) and \( V_{DS} = 5 \text{ V} \). From a comparison of the extracted \( N_{It} \) and \( Q_{ox} \) shown in Fig. 5, we see that the peak values of maximum \( N_{It} \) and \( Q_{ox} \) are about several hundred angstrom away from that of maximum lateral electric field. Furthermore, it was also found that the lateral electric field of stressed device is slightly greater than that of a fresh one since the channel resistance increases due to the localized \( N_{It} \) and \( Q_{ox} \) buildup during the stress. From Fig. 6, the enhancement of electric field after the hot-electron stress can properly explain the increase of substrate current (shown in Fig. 2) since the impact-ionization rate also increases. However, the gate current degradation still can not be explained from the gate current model in [3] and [10]. From the decrease of gate current with stress time as observed in Fig. 2, an additional degradation factor due to oxide damage exists. Therefore, we suggest that the interface states filled with electrons can be served as a new scattering center to the injected hot electron which reduces the hot-electron injection probability.

Fig. 7 is used to support the above argument. In this figure, the measured and simulated gate current characteristics are shown for devices at fresh, after hot-electron stressed, and
Fig. 6. Simulated lateral electric field distributions of devices before and after $I_{G,\text{max}}$ stress.

Fig. 7. Measured and simulated gate current characteristics for devices at fresh, after hot electron stress, and after detrapped conditions. Note that $N_{it}$ is dominant for the $I_{G}$ degradations.

after detrapped conditions at $V_{DS} = 5$ V, respectively. The detrapping step is used to eliminate $Q_{ox}$. The GIDL current is given as a monitor for the oxide charge detrapping, as shown in Fig. 4(b). From the simulation data, we see that the difference of $I_{G}$ between the HE stressed (solid triangles) and the detrapped (solid rectangles) conditions is due to the $Q_{ox}$ effect. Also, the difference between the fresh (solid circles) and the detrapped (solid rectangles) curves is due to the $N_{it}$ effect. Obviously, the $I_{G}$ degradation due to $N_{it}$ is much larger than that due to $Q_{ox}$. This can also explain why the gate current degradation after $I_{G,\text{max}}$ stress is much more serious than that after $I_{G,\text{max}}$ stress. From measurement data, we also observed that the $I_{G}$ degradation rate ($\Delta I_{G}/I_{G}$) first increases gradually with the increasing gate voltage and then decreases. The reasons are as follows. The scattering of injected hot electron due to $N_{it}$ is dependent on the amount of electrons filled in the interface states. As the gate voltage increases, the amount of electrons filled in these states also increase. As a result, we see that $\Delta I_{G}/I_{G}$ increases as the gate voltage increases initially. However, by further increasing the applied gate voltages, the charge screening effect occurs and $\Delta I_{G}/I_{G}$ becomes moderate. From the simulated data shown in Fig. 7, the degradation of $I_{G}$ has been successfully simulated by using the new model described in (4). The parameter, $\alpha_{\text{eik}}$, with $1.16 \times 10^{-12}$ cm$^{-2}$ is used. Furthermore, Fig. 8 shows the measured and simulated gate current characteristics among different stress times. Reasonable agreement of the gate currents between simulated and the measured results has been achieved.

Fig. 9 shows the simulated gate current differences at $V_{DS} = 4$ V and $V_{GS} = 4\ell$ V by considering $N_{it}$ and/or $Q_{ox}$, and their comparison with the measurement data for different stress time. Stress condition is also at $V_{GS} = 5.5$ V and $V_{TS} = 5$ V. It was found that the degradation of gate currents due to $Q_{ox}$ will saturate after a long stress time. Furthermore, it was also seen that degradation of $I_{G}$ is mainly due to the generated $N_{it}$. Finally, as the device is continued to scale, we believe that the effects of $N_{it}$ will become more crucial to the gate current rather than $Q_{ox}$.

VI. CONCLUSION

In summary, we have successfully developed a new gate current model which takes the hot-electron stress generated interface states and oxide trapped charges into account. In this model, we suggest that the interface states will serve as a new scattering center to the injected hot electron and then reduce
the hot-electron injection probability. The generated oxide charges are also introduced as an additional factor affecting the injection potential barrier at the Si–SiO₂ interface. However, the oxide charges are normally located inside the gate oxide such that they have only minor effect to the barrier height. As a result, the influence of \( Q_{ox} \) on the gate current degradation is smaller than that of \( N_c \). In other words, the interface state generation will dominate the gate current degradation not only at the maximum substrate stress condition but also at the maximum gate current stress condition. Results of the simulated gate currents show a good agreement with the measurement data for pre-stressed and post-stressed devices. Finally, the developed new gate current degradation model is especially useful for the investigation of hot carrier reliabilities in EPROM and Flash EEPROM devices, in which the maximum gate current bias is the programming bias and is crucial for these devices. In other words, the proposed gate current degradation model is most useful for studying nonvolatile memory reliabilities.

REFERENCES


