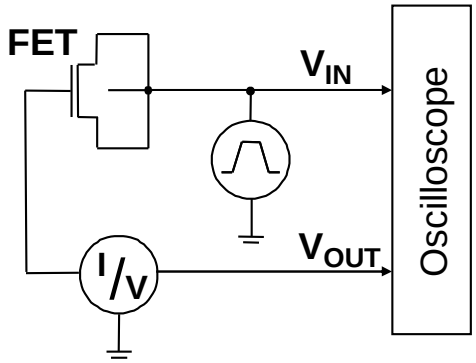

Figure 5.6: Schematic setup used to obtain pulsed C-V characteristics. The voltage pulse applied to the source, drain and substrate (V_{IN}) is recorded using a digital oscilloscope together with the output voltage (V_{OUT}) of the current−voltage amplifier.

technique is shown in Fig. 5.8.

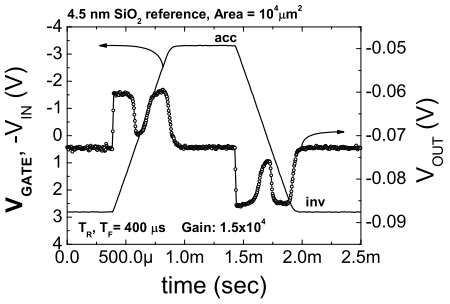

Figure 5.7: Voltage traces in a pulsed C-V measurement recorded with a digital storage oscilloscope.

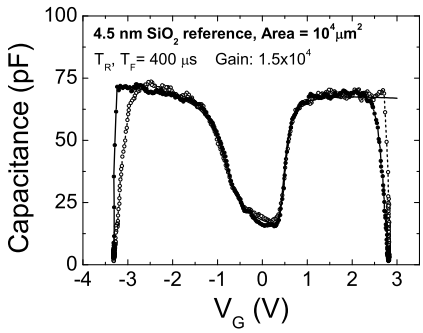

Figure 5.8: Comparison of the C-V characteristic measured using the pulsed C-V and conventional High Frequency technique.

As can be seen a good quantitative agreement is obtained for the sample with a conventional gate dielectric. The small deviations at the end of the sweep range are caused by the non-ideal shape of the linear voltage pulse provided by the pulse generator. In principle, this effect can be taken into account by detailed data analysis.

5.2.2 C-time traces

Another way to monitor transient effects in MOS structures is based on measuring capacitance-time traces in accumulation. In this case the LCR meter is used in a sampling mode and the capacitance reading is recorded as a function of time. The ability to record transient effects strongly depends on the resolution of the LCR meter and the C-V swing from depletion to accumulation. The LCR meter is well known for its accuracy and even small changes in the capacitance can be measured

reliably. The C-V swing (dC/dV), however, is large around flatband and strongly decays when sweeping the device into strong accumulation. Furthermore, the swing depends on the EOT, thinner dielectrics show a more pronounced swing compared to thicker ones.

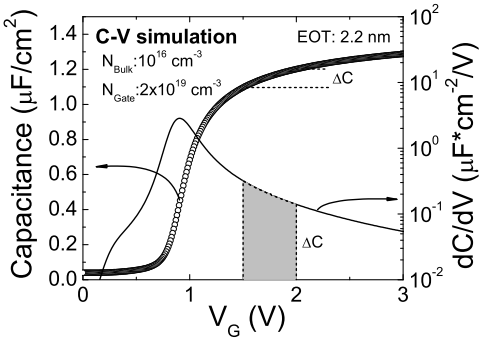

Figure 5.9: Simulated C-V characteristic for a MOS capacitor on n-type Si substrate. In addition, dC/dV is shown to illustrate the capacitance change as function of gate bias.

In summary, the capacitance-time trace presents another possibility to monitor instabilities in thin high-ε gate dielectrics where effects already occur in weak accumulation.

5.2.3 Pulsed I_D-V_G measurements

The pulsed I_D-V_G measurement is another viable technique to study the trapping phenomena. The major advantage of this technique over the pulsed C-V method is the reduced sensitivity to the parasitic gate leakage current. Unless the I_D-V_G characteristic is distorted by the leakage current, this procedure can be used to monitor V_T shifts in the μs time range.

The schematic measurement setup for the pulsed I_D-V_G measurement is given in Fig. 5.10. In this setup, the MOSFET is used in an inverter circuit with a resistive load (R_L). A small DC bias is applied to the resistor which acts together with the channel resistor as a voltage divider. The I_D-V_G characteristic is obtained by applying a trapezoidal (triangular) pulse to the gate and recording the drain voltage using a digital oscilloscope. From the measured voltage traces the I_D-V_G characteristic can be constructed using the following relationship

$$I_D = \frac{100mV}{V_D}\left(\frac{100mV - V_D}{R_L}\right) \tag{5.2}$$

where V_D is the measured drain voltage and R_L the resistive load of the inverter [48]. Due to the use of a voltage divider the drain voltage dynamically changes during the measurement. However, this effect can simply be eliminated by normalizing the extracted drain current to a constant drain voltage, which is given by the term $100m\,\mathrm{V}/\mathrm{V}_D$ in Eq. 5.2. It is noted that this normalization is only correct, when the MOSFET is operated in the linear regime, which limits the range for the DC bias applied to the resistor.

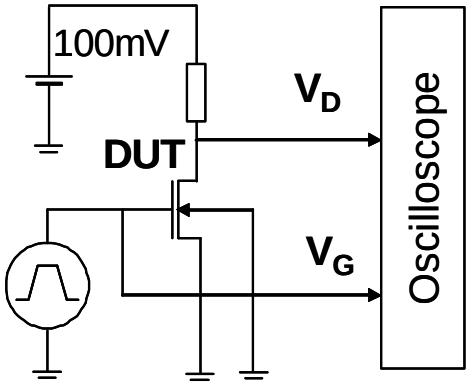

Figure 5.10: Schematic of setup used for pulsed I_D-V_G measurements. The FET is used in an inverter circuit with a resistive load (R_L). From the voltage - time traces (V_G(t), V_D(t)) the I_D-V_G characteristic is extracted.

When replacing the resistor with a current-voltage amplifier connected to the source terminal this problem can be circumvented and a constant bias can directly be applied to the drain terminal of the device.

5.2.4 Drain current degradation (I_D-time)

Another way of looking at the stability of MOSFETs is monitoring the drain current as a function of time, which is basically a "sense at stress" procedure. When combining the conventional techniques with the pulsed measurement setup the time base can be extended by several decades, going from μs to >100 ks. Moreover fast degradation in the drain current can be captured by the combined technique.

The instability in the drain current can either result from a loss / gain in the inversion charge due charge trapping effects or from an enhancement in the scattering mechanism due to generation of scattering centers like interface states. In any case, the stability of the device parameters need to be guaranteed for a specified lifetime of the circuit, which is commonly 10 years.

5.2.5 Charge pumping technique

The charge pumping (C-P) technique was already briefly discussed in Section 2.2.3 and the basics given according to conventional SiO_2 gate dielectrics. Based on the charge pumping theory [14] defect states at the interface are being charged with inversion carriers which subsequently recombine with majority carriers in the Si substrate. The amount of carriers recombining is determined by the pulse shape (rise and fall time). Due to the fact that the bulk defect density in conventional SiO_2 is very small, no significant contribution from bulk traps to the charge pumping current measured at the Si substrate is to be expected. If a typical initial trap density of $10^{15} - 10^{16} cm^{-3}$ [49] is used for thermally grown gate dielectrics, the contribution from bulk oxide traps is estimated to be in the order of $\sim 10^9 cm^{-2}$. Typical interface trap densities in MOSFETs with thermally grown gate dielectrics are frequently reported in the $10^{10} cm^{-2}$ range. Due to the low density of bulk traps in conventional gate dielectrics their contribution to the charge-pumping current cannot easily be separated.

In the case of high-ε gate dielectrics, due to the presence of a high bulk defect density, it will be demonstrated that charge pumping can be used to sense these defect states.

In Fig. 5.11 the basic charge pumping modes are shown schematically. In a conventional charge pumping measurement the amplitude of the gate pulse is fixed and the base level is swept from either accumulation to inversion or vice versa. The source and drain terminals are connected together and usually put to ground potential, whereas the charge pumping current is measured at the substrate terminal. In principle, when the gate leakage is negligible the same current should also be measured at the source plus drain terminals.

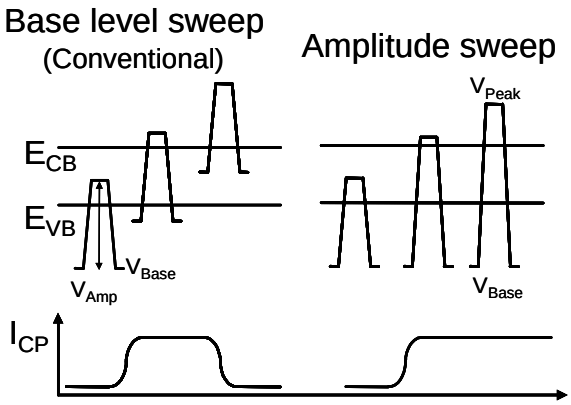

Figure 5.11: Schematic of charge pumping measurement using the conventional base level sweep and an amplitude sweep. The charge pumping current (I_{CP}) measured at the Si substrate is shown for comparison.

An alternative charge pumping procedure is given when the base level of the pulse is kept constant and the amplitude is varied. This procedure is called amplitude sweep charge pumping. In case of conventional gate dielectrics both techniques yield similar results, whereas for high-ε gate dielectrics significant differences can be observed.

To understand the cause of the discrepancy between the two charge pumping modes we first discuss the different current contributions in a charge pumping measurement as illustrated in Fig. 5.12.

Beside the contribution from recombination of interface states, gate leakage is known as a source of additional substrate current. Especially in ultra-thin gate dielectrics the contribution exponentially increases with decreasing oxide thickness. Therefore, experimental procedures have been proposed which can efficiently suppress this effect [50]. In addition to interface states also charging and discharging of bulk defect states in the high-ε gate dielectrics contributes to the substrate current. Other than for conventional gate dielectrics this component can add significantly to the substrate current, due to the presence of a high trap density in the high-ε material. In particular a strong voltage and frequency dependence is expected due to the spatial and energy distribution of the bulk defects.

Furthermore, the recombination of inversion carriers with bulk majority carriers can also contribute to the substrate current. This effect is known in the literature as the geometrical component [17]. Therefore, short channel devices ($< 1\,\mu$m) are

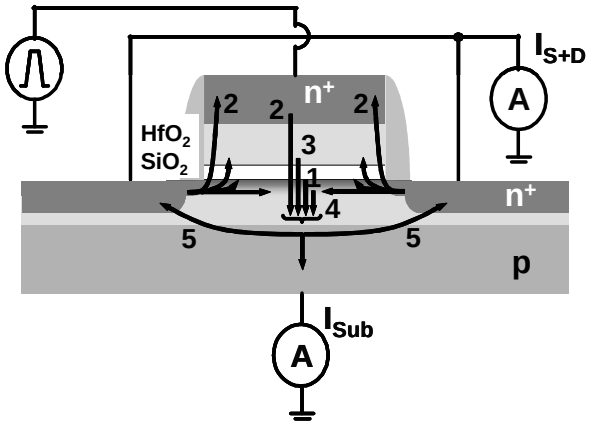

Figure 5.12: Possible current contributions in a charge pumping measurement with alternative gate dielectrics. Beside the recombination current due to interface states (1), the gate current contribution (2), charging and discharging of bulk defects (3), recombination of inversion carriers (4) and minority carrier diffusion (5) due to electron injection from the HfO_2 layer need to be considered.

commonly used for charge pumping measurements, where this effect is minimized. When considering the charging and discharging process during a charge pumping cycle in more detail, minority carriers injected from the inversion layer get emitted when majority carriers in the Si substrate are accumulated again. In order to contribute to the charge pumping current these carriers have to recombine in the substrate. However, the minority carriers in the Si have a certain lifetime and can diffuse back to the source and drain junctions. Therefore, only a fraction of the emitted carriers will effectively contribute to the substrate current. This effect is expected to significantly depend on the device geometry.

6

Charge trapping and V_T-instabilities in alternative gate dielectrics

One of the major integration challenges for high-ε gate dielectrics is the electrical stability of MOS devices. In order to correctly assess the stability of such devices several measurement techniques have been introduced in the previous section. These techniques are applied to SiO_2 / Al_2O_3 gate stacks with TiN electrodes as well as for SiO_2 / HfO_2 with conventional n-type poly Si gate electrodes. Both material systems will be discussed in the following separately.

6.1 Gate stacks containing Al_2O_3

The details concerning the device fabrication process can be found in Section 3. The experimental details for the Al_2O_3 gate stacks can be summarized as follows. First, a thin ($\sim$ 1 nm) thermally grown SiO_2 layer was formed by Rapid Thermal Oxidation (RTO) followed by the deposition of 7, 15, and 20 nm Al_2O_3. Samples with a 5 minute PDA at 700, 800 and 900°C in N_2 are compared to samples without post deposition treatment. PVD TiN was used as gate electrode in the non-self aligned gate last process as described in Section 3.2. All samples have received a final passivation anneal in N_2/H_2 at 420°C for 30 minutes.

6.1.1 I_G-V_G method as monitor for charge trapping

The I-V characteristic is a useful monitor for charge trapping only in thick gate stacks containing high-ε gate dielectrics. The gate I-V of such stacks, where FN-like tunneling is the dominant conduction mechanism, is sensitive to the amount, nature (positive or negative) and location of the trapped charge. Furthermore, from sequential I-V traces information on the stability of the trapped charge can be obtained.

As can be seen in Fig. 6.1 the gate I-V is significantly steeper on the reverse trace indicating a build up of negative charge. Whereas for negative polarity the reverse trace increases compared to the forward trace suggesting a build up of positive charge in this case. Furthermore, it seems that charge trapping is reversible to a large extent when the polarity is changed. Only minor differences are seen

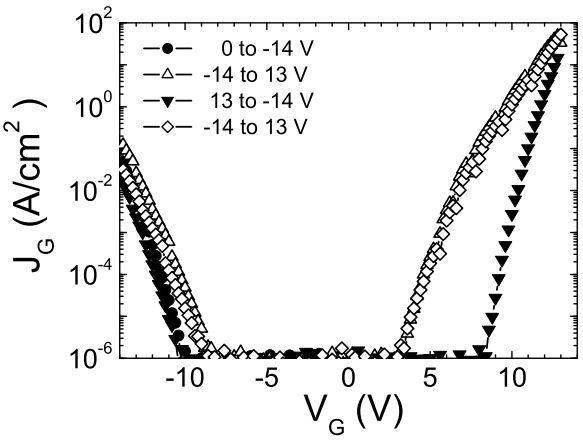

Figure 6.1: Hysteresis in the I-V characteristics of an *as deposited* 15 nm Al_2O_3 layer on n-type Si substrate. The observed asymmetry in the I-V hysteresis is attributed to trapping of negative charge at positive bias and positive charge for negative polarity, respectively.

between the initial and second hysteresis trace which shows that the total amount of traps remains unchanged by performing the sweep itself.

To study the stability of the trapped charge in the dielectric after removing the gate bias, multiple I_G-V_G traces are measured (Figs. 6.2 and 6.3). Here a 15 nm Al_2O_3 layer with and without PDA on n-type Si substrates with TiN electrodes was used. The I-V traces were taken from $0V$ to V_{G_max}, with V_{G_max} increasing from $+5V$ to $+12V$ in steps of $1V$. For positive gate bias (substrate injection) negative charge trapping is observed in all cases. A build up of negative charge in the gate stack causes a positive I-V shift. Therefrom a bulk defect density in the Al_2O_3 can be estimated.

For *as deposited* Al_2O_3 (Fig. 6.2), an I-V shift of $\sim 2V$ is observed for $V_{G\_max} = 12V$, which correlates to a charge areal density of $\sim 1 \cdot 10^{13} cm^{-2}$. Assuming bulk charging of the Al_2O_3 layer a bulk volume defect density of $\sim 6 \cdot 10^{18} cm^{-3}$ [51] is extracted.

When comparing the multiple I-V traces of *as deposited* (Fig. 6.2) and high-temperature annealed Al_2O_3 (Fig. 6.3) it is evident that the charge trapping after the high-temperature PDA radically changed. The anneal in N_2 at 900°C apparently introduced new types of defects, which are charged at low gate bias and shallow defects which cause transient charging effects. The transient charging effects result in the partial recovery after the gate bias is taken off. Therefore reduced I-V shifts are measured for subsequent traces. The presence of transient charging effects makes an assessment of charge trapping more complicated.

The charge trapping behavior of 7 and 20 nm thick *as deposited* Al_2O_3 layers based on sequential I-V traces is similar to the data shown in Fig. 6.2. Comparable results to those in Fig. 6.3 were also obtained by varying the PDA temperature, indicating that 700°C and 800°C anneals in N_2 result in similar formation of new types of defects as concluded from Fig. 6.3.

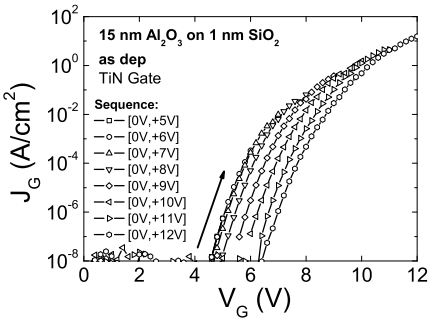

Figure 6.2: Multiple I-V traces taken on a 15 nm *as deposited* Al_2O_3 layer using the measurement sequence as indicated in the figure. The charge trapped during the I-V sweep remains stable at $0V$ gate bias.

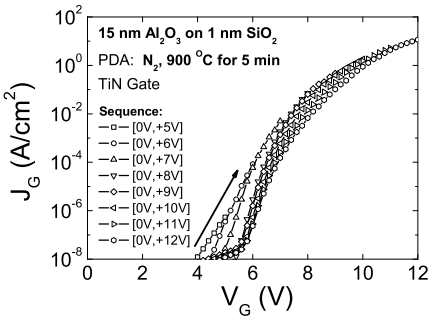

Figure 6.3: Multiple I-V traces taken on a 15 nm high-temperature annealed Al_2O_3 layer using the measurement sequence indicated in the figure. Detrapping at low gate bias results in partial charge recovery between subsequent traces.

6.1.2 C-V shifts and I-t traces as monitor for charge trapping

An alternative technique to monitor charge trapping in the gate dielectrics is given by the C-V measurement. A shift in the C-V characteristic is attributed to the build up of charge in the dielectric and can be calculated using Eq. 6.1

$$\Delta V = \int_0^{T_{OX}} \frac{1}{\varepsilon_0 \cdot \varepsilon_{OX}(x)} (T_{OX} - x)\rho_{OX}(x)dx \tag{6.1}$$

where T_{OX} is the dielectric thickness, $\rho_{OX}(x)$ the distribution of trapped charges given in $[Ccm^{-3}]$, ε_{OX} the dielectric constant and x the distance from the Si substrate. The effective trapped charge density $Q_{OX}[Ccm^{-2}]$ is defined as

$$\Delta V = \frac{1}{C_{OX}} Q_{OX} \tag{6.2}$$

where C_{OX} is the oxide capacitance in $[Fcm^{-2}]$. It corresponds to a charge sheet located at the interface between the Si and the dielectric. The effective trapped charge density is for convenience converted to ΔN_{OT} which is the equivalent charge trap density given in $[cm^{-2}]$.

A typical example for multiple C-V traces taken on a 15 nm *annealed* Al_2O_3 layer on n-type Si substrates is shown in Fig. 6.4. The C-V curves were taken starting with a trace from 0 to $+5V$ followed by retracing from $+5V$ to 0. Afterwards a "trace - retrace - trace" sequence from 0 to $+6V$ were taken. As can be seen, the reverse sweep after a maximum bias of $+5V$ is significantly shifted towards positive values due to the build up of negative charge. However, when the subsequent trace is taken some fraction of the charge gets detrapped, leading to a recovery of the C-V shift. The subsequent trace was taken to $+6V$, which resulted in a more pronounced shift and recovery as well. Based on Eq. 6.2, for a $\Delta V \sim 1V$ the equivalent charge trap density is $\sim 3.5 \cdot 10^{12} cm^{-2}$.

The transient discharging observed in *annealed* Al_2O_3 is not found in the *as deposited* layers. This is illustrated by comparing Fig. 6.5 with Fig. 6.6. In Fig. 6.5 both the gate current during CVS and the corresponding voltage shift from the C-V characteristic are shown for a 15 nm *as deposited* Al_2O_3 layer. As can be

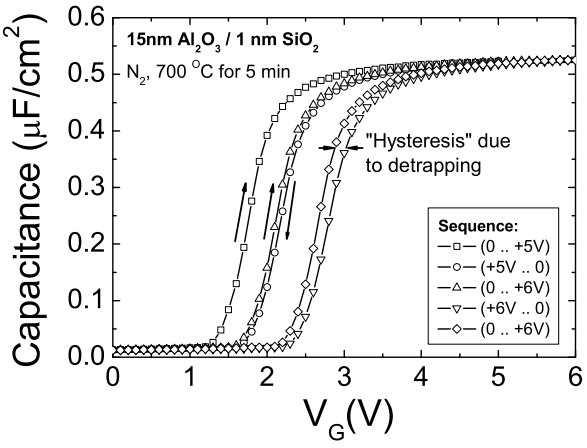

Figure 6.4: Multiple HF C-V traces taken on a 15 nm *annealed* Al_2O_3 layer using the sequence as indicated in the figure. From the C-V shift a stable and transient charging component is evident.

seen, the gate current versus injected charge continuously decays due to the build up of negative charge. Moreover, when the CVS is interrupted to measure the C-V characteristic the gate current does not significantly recover, which indicates little detrapping during sensing. When the same measurement procedure is applied to a 15 nm *as deposited* Al_2O_3 layer, as shown in Fig. 6.6, it is evident that for the same amount of injected charge larger shifts are observed and furthermore significant detrapping occurs during sensing. Detrapping during the sense measurement results in a spike in the current charge trace, due to the recovery of the gate current.

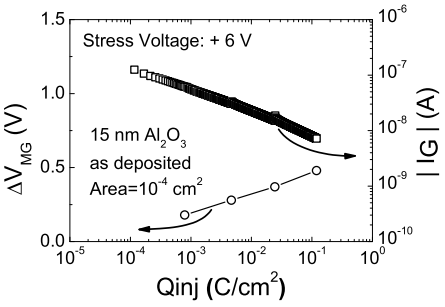

Figure 6.5: Gate current versus injected charge for a 15 nm *as deposited* Al_2O_3 layer using a CVS as indicated in the figure. The stress was interrupted on a logarithmic time scale to monitor the C-V shift. No significant charge loss was observed for the *as deposited* layer.

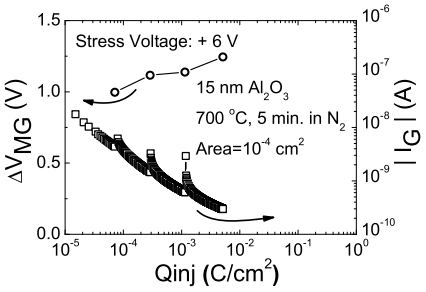

Figure 6.6: Gate current versus injected charge for a 15 nm *annealed* Al_2O_3 layer using a CVS as indicated in the figure. In contrast to Fig. 6.5, significant charge loss during sensing results in spikes in the current - charge trace.

A more detailed analysis of the I-V and C-V shift was carried out on a 15 nm *as deposited* Al_2O_3 layer, where detrapping was found to be negligible. The gate I-V

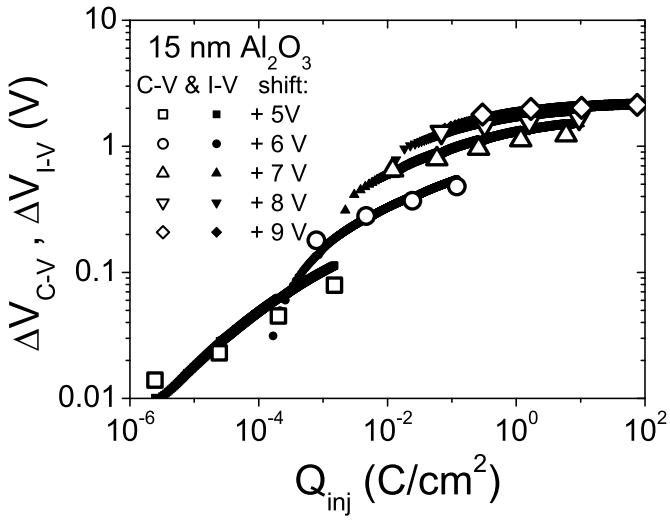

Figure 6.7: Comparison of I-V and C-V shift measured on a 15 nm *as deposited* Al_2O_3 layer. Good agreement between both I-V and C-V shift indicates that the trapped charge is distributed throughout the dielectric.

shifts were extracted from current time traces as shown in Fig. 6.5 using

$$\Delta V_{I-V} = \frac{\Delta I}{d(\log(I))/dV} \tag{6.3}$$

where $d(\log(I))/dV$ is the slope of the reconstructed I-V characteristic of the charge free stack. This procedure can only be used in the absence of transient charge loss and hence cannot be applied to annealed stacks. As can be seen in Fig. 6.7, identical shifts are obtained from the C-V and I-V characteristics. This indicates that a large fraction of the trapped charge is located outside the tunneling distance and is likely trapped in the bulk of the Al_2O_3 layer, as will be discussed in Section 6.1.4.

If we ignore the presence of transient charge loss for *annealed* Al_2O_3 layer for a moment and apply the conventional "stress and sense" procedure as discussed in Section 5.1.3, the charge trapping behavior of *as deposited* and *annealed* layers can be compared to first order. Fig. 6.8 shows the result. As mentioned previously, larger voltage shifts are observed for the *annealed* stacks independent of the stress voltage applied. The effect is more pronounced at low injected charge densities, again suggesting the pre-existence of defects which are efficiently charged at a low fluence. For large injected charge densities the differences between the *as deposited* and the *annealed* layer with respect to stable trapped charge become less significant. In both cases equivalent charge densities of $> 5 \cdot 10^{12} cm^{-2}$ are measured.

When comparing the measured charge build up with a first order trapping model as discussed in Section 2.3.1 it becomes clear that such a simplified model cannot explain the observations. A parameterization of the data would likely require the use of multiple capture cross sections and the inclusion of field dependent detrapping. Without the latter the increase in the saturation charge density (leveling off, plateau) with increasing gate voltage or better with increasing current density cannot be explained [52].

The trapping behavior for negative gate bias (gate injection) is radically different compared to positive bias. For gate injection positive charge (hole) trapping is

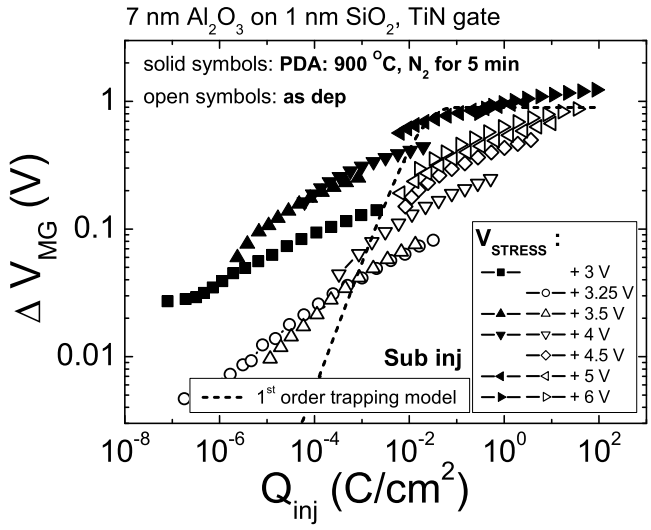

Figure 6.8: Comparison of the C-V shifts between the *as deposited* and an *annealed* Al_2O_3 layer. Larger shifts are observed for the *annealed* sample compared to the *as deposited* layer.

dominant in all Al_2O_3 cases studied, resulting in negative ΔV_{MG} as shown in Figs. 6.9 and 6.10. For *as deposited* stacks (Fig. 6.9) hole trapping shows no significant dependence on the gate bias, it only seems to depend on the amount of injected charge. No indication of saturation is observed up to dielectric breakdown of the gate stack. Furthermore, when the device is biased in inversion (positive gate bias) during the C-V sensing, a large fraction of trapped holes recombine, as shown in Fig. 6.9. This effect is also seen in single layer SiO_2 stacks, and it was attributed to the recombination of trapped holes near the interface by tunneling [19]. In the SiO_2 / Al_2O_3 stack, hole trapping predominantly occurs in the interfacial SiO_2 layer as discussed in more detail in Section 6.1.4. In addition to hole trapping a rapid build up of interface states is observed for gate injection, indicating degradation of the interfacial SiO_2 layer, which is not observed for substrate injection [53].

In high temperature annealed Al_2O_3 films, hole trapping is also the dominant mechanism for stress under negative gate bias (Fig. 6.10). The trapping kinetics of annealed stacks, however, significantly differ from that observed in *as deposited* films, showing that the thermal history strongly influences the electrical behavior.

Positive charge trapping in *as deposited* Al_2O_3 films is also confirmed by the current charge traces shown in Fig. 6.11. The build up of positive charge during CVS leads to a current increase. For the *annealed* stacks, however, hole trapping as it is sensed by the C-V measurement is not always evident from the current time traces. As can be seen in Fig. 6.11, at low negative gate bias the current decreases as a function of the injected charge. This indicates that the gate current is sensing electron trapping, which apparently contradicts the results of the C-V measurements. Indeed, in order to explain the current reduction together with a negative ΔV_{MG} electron and hole trapping have to occur simultaneously. The C-V is sensing the charge from the Si substrate and therefore is more sensitive to the trapped charge close to the substrate. This indicates that hole trapping occurs either in the interfacial SiO_2 layer or near the interface between the two dielectrics. At the same time, electron trapping takes place in the bulk of the Al_2O_3 layer,

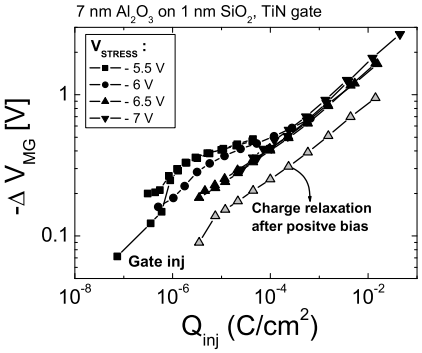

Figure 6.9: C-V shifts measured on a 7 nm *as deposited* Al_2O_3 layer on p-type Si substrate. Negative voltage shift indicates build up of positive charge. No saturation of positive charge trapping is observed up to dielectric breakdown.

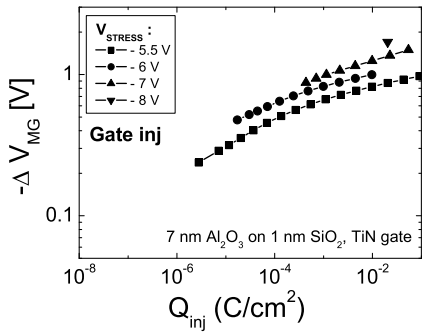

Figure 6.10: C-V shifts measured on a 7 nm *annealed* Al_2O_3 layer on p-type Si substrate. Positive charge trapping results in negative voltage shift, however, the trapping kinetics are different compared to Fig. 6.9.

which is sensed by the gate I-V measurements.

For substrate injection the current time traces confirm the build up of negative charge as sensed by the C-V measurement (shown in Fig. 6.12), which results in a significant current reduction. Furthermore, the transient charging effects are again confirmed in the *annealed* stacks, while they are not present in the *as deposited* films. The transient charge loss and therewith the observed spikes in the current charge traces are consistent with the sequential I-V traces shown for *annealed* stacks, as discussed in the context of Figs. 6.3 and 6.6.

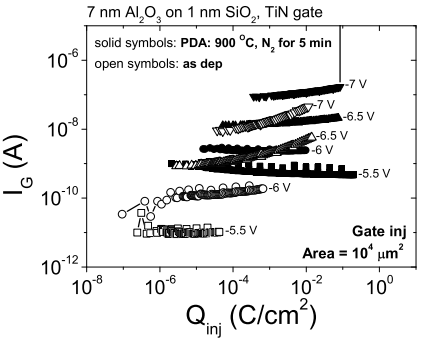

Figure 6.11: Comparison of gate current versus injected charge for a 7 nm *as deposited* and *annealed* Al_2O_3 layer. Charge injection for the gate electrode using CVS at voltages as labeled in the figure.

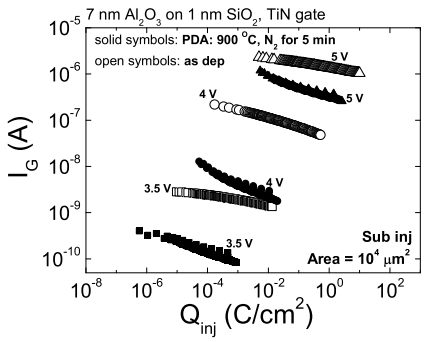

Figure 6.12: Comparison of gate current versus injected charge for a 7 nm *as deposited* and *annealed* Al_2O_3 layer. Charge injection from the n-type Si substrate using CVS at voltages as labeled in the figure.

In summary, charge trapping in SiO_2 / Al_2O_3 gate stacks strongly depends on the injection polarity and post deposition treatment. For substrate injection electron trapping of pre-existing defects is dominant, whereas hole trapping occurs for gate injection. High temperature annealing in N_2 introduced new types of

defects, which result in transient charging effects. These transient effects cannot be captured by conventional measurement techniques. Therefore novel (time-resolved) measurement techniques down to the μs range are required, as it was discussed in Section 5.2.

All these experiments show that the electrical instability arises from filling of pre-existing defects in the gate stack, with a density in the order of $\sim 10^{18}$ to $10^{19} cm^{-3}$, which is 2-3 orders of magnitude higher than in conventional SiO_2 gate dielectrics [49]. Therefore charge trapping is expected to remain strong also in scaled stacks, which was directly confirmed by results on 3 nm Al_2O_3 films with a sub 2 nm EOT [54].

6.1.3 Interface state generation

The properties of the interface between the Si substrate and the gate dielectric are crucial parameters for the MOSFET performance. Therefore, the Si / SiO_2 interface has been investigated intensively since the early stage of MOSFET fabrication. The high density of interfacial defect states inherent due to the mismatch of the Si substrate and the thermally grown SiO_2 can be passivated efficiently by using a standard N_2 / H_2 passivation anneal. Values typically reported for a conventional SiO_2 gate stacks are in the low $10^{10} cm^{-2}$ range. However, when applying an electrical stress to the MOS device, the formation of interface states is commonly observed and is attributed to the release of hydrogen at the Si / SiO_2 interface, independent of the stress polarity.

Achieving good interface properties is also considered an important factor for the integration of high-ε gate dielectrics into future CMOS technologies. The standard passivation anneal used for SiO_2 is commonly adopted and applied to high-ε dielectric, but does not always yield the desired results. In this section we are not going to address the efficiency of the passivation anneal, but rather discuss the formation of interface states due to electrical stress.

The initial interface state density as measured by the equilibrium controlled quasi static C-V technique, as introduced in Section 2.1.2, was found to be $\sim 5 \cdot 10^{11} cm^{-2}$, which may not be adequate low for good MOSFET performance. However, it is sufficiently low to monitor the increase of interface states as a function of stress polarity as can be seen in Fig. 6.13. Compared to the fresh device, a rapid build up of interface states in combination with positive charge trapping are observed after applying a negative polarity stress already at fluencies as low as $\sim 10^{-3} C cm^{-2}$. In the previous section it was mentioned, that a fraction of the positive charge recombines when the polarity is reversed. As shown in Fig. 6.13, when applying a sufficient positive bias to the gate (where electron injection from the Si substrate begins) the positive trapped charge can completely be compensated by electron trapping, resulting in a net positive voltage shift compared to the initial trace. The minimum capacitance, however, remains unaffected indicating that the formation of interface states is irreversible.

The physical process causing the formation of interface states for gate injection is expected to be similar to that for single layer SiO_2. Electrons injected from the TiN gate into the conduction band of the Al_2O_3 layer are expected to reach the Si substrate with a high energy, because of ballistic transport through the thin SiO_2 layer. It is well known that such hot carriers are the cause for interface state generation in conventional SiO_2 gate stacks. Therefore, in SiO_2 / Al_2O_3 gate stacks the formation of interface states for negative gate bias is most likely of similar origin. The shape of the C-V characteristic after negative bias stress, as shown in Fig. 6.14, for an *as deposited* Al_2O_3 layer is possibly caused by interface states (N_{it}) due to the high hydrogen content in this film. In the literature, similar distortions are

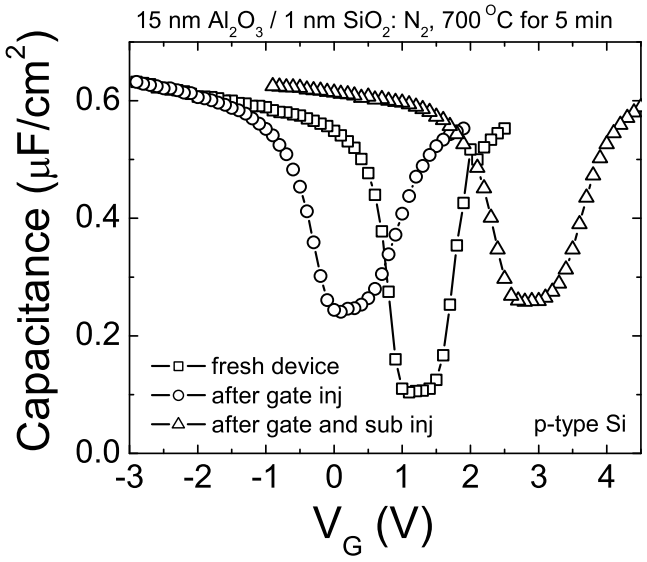

Figure 6.13: Quasi static C-V characteristics of a fresh Al_2O_3 sample annealed in N_2 at 700°C followed by negative and positive polarity stress. Irreversible interface degradation is found after negative bias stress.

reported for Al gated MOS devices which have not seen a high temperature process during device fabrication. In general, *as deposited* layers appear to be more prone to hot carrier degradation.

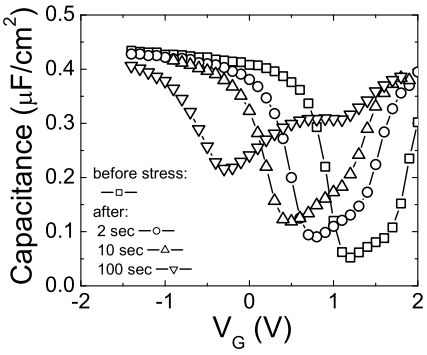

Figure 6.14: Quasi static C-V characteristics of *as deposited* 15 nm Al_2O_3 / 1 nm SiO_2 on a p-type Si substrate before and after CVS at V_G=-11V for 2, 10 and 100 sec.

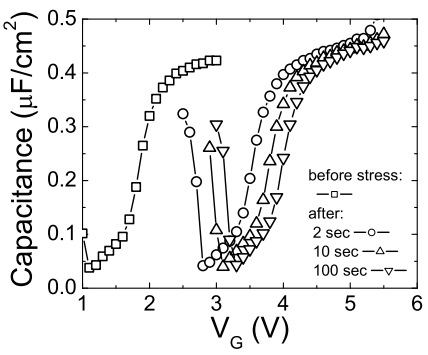

Figure 6.15: Quasi static C-V characteristics of *as deposited* 15 nm Al_2O_3 / 1 nm SiO_2 on a n-type Si substrate before and after CVS at V_G=9V for 2, 10 and 100 sec.

In contrast as shown in Fig. 6.15, for substrate injection no distortion of the C-V characteristic is observed up to a fluence of $\sim 10\,\mathrm{Ccm}^{-2}$. The absence of interface state generation for this polarity case indicates that carrier heating does not occur, which may be prevented by the high density of defects in the Al_2O_3 layer. It is also

possible that intermediaries, such as protons or holes cannot be transported back to the Si substrate.

6.1.4 Location of trapped charge

In the past, gate current-voltage measurements in combination with capacitance voltage measurements have been successfully applied to thick SiO_2 films to quantify the amount of trapped charge and locate its position in the dielectric [55]. The same technique was adopted here to investigate SiO_2 / Al_2O_3 gate stacks.

For charge location analysis, I-V characteristics and C-V characteristics were taken from neighboring devices with Al_2O_3 gate stacks before and after constant voltage stress for 1, 10 and 100 sec. In Fig. 6.16 the corresponding I-V traces are shown. As can be seen, positive I-V shifts are observed for both injection polarities indicating positive charge trapping for negative gate bias and negative charge trapping for positive gate bias, which is consistent with the data presented in Section 6.1.1 and 6.1.2 and the results obtained by C-V sensing as shown in Figs. 6.14 and 6.15. Considering the sign of the trapped charge I-V and C-V sensing provide the same qualitative information.

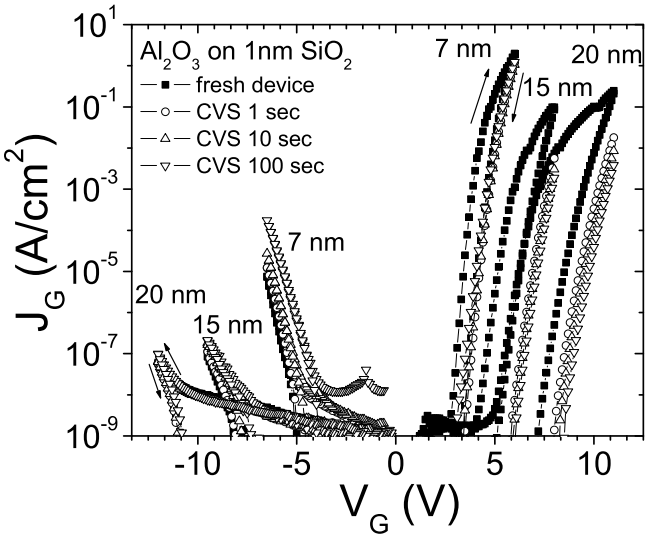

Figure 6.16: I-V characteristic of SiO_2 / Al_2O_3 stacks before and after CVS stress at negative and positive gate bias, respectively. Positive shift in I-V characteristic is obtained independent of the injection polarity.

To enable a more quantitative analysis of the charge trapping data, I-V and C-V shifts have been extracted from data as shown in Figs. 6.14 and 6.15 and are plotted as a function of injected charge for substrate and gate injection in Figs. 6.17 and 6.18, respectively. Data for 7, 15 and 20 nm Al_2O_3 are compared. It is important to note that the gate I-V is sensing the trapped charge at the injecting side, which means from the gate electrode for gate injection and from the Si substrate for substrate injection. By combining I-V and C-V sensing for gate injection the centroid of the trapped charge can be determined from the measured

voltage shifts [55]. For substrate injection this technique is not applicable, because I-V and C-V are sensing the trapped charge from the Si substrate side. However, from the absolute magnitude of the measured voltage shifts and the dependence on the Al_2O_3 thickness information on the charge location can be derived for this polarity as well.
In the following the results for substrate and gate injection are discussed separately:

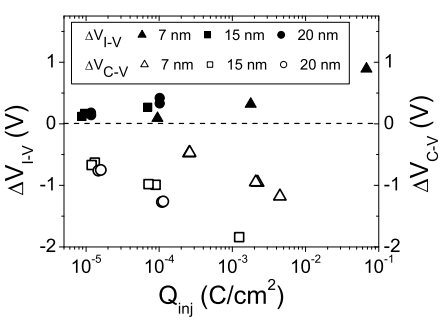

Figure 6.17: C-V and I-V shifts for *as deposited* Al_2O_3 / SiO_2 stacks on p-type Si substrates due to CVS at negative gate bias.

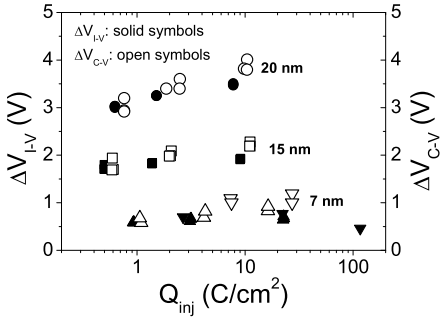

Figure 6.18: C-V and I-V shifts for *as deposited* Al_2O_3 / SiO_2 stacks on n-type Si substrates due to CVS at positive gate bias.

Gate Injection: According to the literature [55] the centroid of the positive charge can be extracted as follows,

$$\Delta V_{I-V} = \frac{Q_{I-V}}{C_{OX}}(1-\overline{x}) \tag{6.4}$$

and

$$\Delta V_{C-V} = \frac{Q_{C-V}}{C_{OX}}\overline{x} \tag{6.5}$$

where Q_{I-V} and Q_{C-V} are the injected charges during CVS, ΔV_{I-V} and ΔV_{C-V} are the corresponding voltage shifts for the gate I-V and C-V, respectively, and $\overline{x} = x/EOT$ is the normalized distance from the gate electrode. The centroid of the positive charge can now be calculated using Eq. 6.4 and 6.5.

$$\overline{x} = \frac{1}{1+\frac{|\Delta V_{I-V}|}{|\Delta V_{C-V}|}} \tag{6.6}$$

The value for the charge centroid, $1-\overline{x}$ and its absolute 'electrical' distance from the Si substrate, $(1-\overline{x})\cdot EOT$, calculated from the data in Fig. 6.17, are summarized in Table 6.1. The physical distance of the positive charge from the SiO_2 / Al_2O_3 interface (d_{int}) can be calculated as follows,

$$d_{int} = \frac{\varepsilon_{Al_2O_3}}{\varepsilon_{SiO_2}}\Big\{\big[(1-\overline{x})\cdot EOT\big] - t_{SiO_2}\Big\} \tag{6.7}$$

As can be seen in Table 6.1, the positive charge is located within 0.26 - 3.2 nm from the interface between the two dielectrics, weakly dependent on the Al_2O_3 thickness and the applied stress condition. This demonstrates that during CVS positive charge trapping near the SiO_2 / Al_2O_3 interface is the origin of flatband and gate I-V shifts for gate injection.

Table 6.1: Centroid of positive charge calculated from data in Fig. 6.17 using Eq. 6.4 to 6.7.

Physical thickness (nm)	EOT (nm)	C-V shift (V)	I-V shift (V)	$1-\overline{x}$	$(1-\overline{x})\cdot$ EOT (nm)	$d_{int}*$ (nm)
7	4.3	-0.95	0.32	0.25	1.1	0.26
15	7.4	-0.65	0.15	0.19	1.4	1.06
		-1.00	0.27	0.21	1.6	1.6
20	9.4	-0.75	0.15	0.17	1.6	1.6
		-1.25	0.38	0.23	2.2	3.2

$* \; t_{SiO_2} = 1nm \qquad \varepsilon_{Al_2O_3} = 10.3$ [56]

Substrate Injection: Due to the fact that I-V and C-V shifts are both measured from the Si substrate side the method used previously cannot be applied. However, qualitative information on the charge location can be obtained, because similar voltage shifts are observed for C-V and I-V measurements, independent of the Al_2O_3 thickness. This is only possible if the charge is located outside the tunneling distance.

In MOS devices with conventional SiO_2, the tunneling distance t_{TUNNEL} under Fowler-Nordheim injection can be estimated using $t_{TUNNEL} = E_{OX} \cdot \Phi_B$, where E_{OX} is the oxide field and Φ_B the barrier height, yielding values of 5 to 2.5 nm for an E_{OX} ranging from 6 to 12 MV/cm. These values exceed the interfacial SiO_2 thickness in the dual layer stacks, suggesting that the negative charge is located in the bulk of the Al_2O_3 layer. From the measured voltage shifts (1V to 4V for Al_2O_3 layers ranging from 7 to 20 nm (see Fig. 6.18)), a bulk charge density in the order of $2-4 \cdot 10^{18} cm^{-3}$ in as deposited Al_2O_3 layers is calculated.

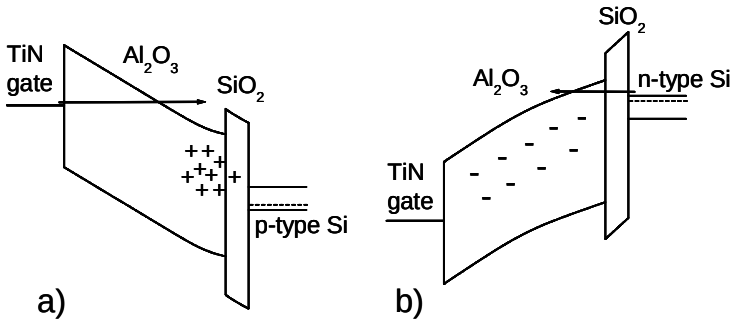

Figure 6.19: Schematic drawing of the energy band diagram of an SiO_2 / Al_2O_3 stack after CVS stress at negative (a) and positive (b) gate bias. Location and nature of trapped charge was obtained by combining C-V with I-V sensing.

6.1.5 Summary

Due to the asymmetry of the SiO_2 / Al_2O_3 gate stack and the difference in work function between the TiN electrode and the Si substrate strong asymmetric charge trapping effects are observed, as summarized in Fig. 6.19. Positive charge trapping near the SiO_2 / Al_2O_3 interface dominates for gate injection. For substrate injection, negative charge trapping in the bulk of the Al_2O_3 layer is observed. In addition, significant interface state generation was measured only for gate injection. It is

possible that strong carrier heating occurs across the interfacial SiO_2 layer leading to oxide degradation as in conventional SiO_2 gate stacks. For substrate injection electron transport may be defect limited preventing heating and degradation of the interfacial layer [53].

6.2 Gate stacks containing HfO_2

Charge trapping in HfO_2 containing gate stacks was investigated using the conventional hysteresis techniques together with the time resolved measurement techniques as introduced in Section 5.2. Therefore, n-channel MOSFETs were fabricated using a self aligned transistor flow as described in Section 3.3. HfO_2 layers were deposited by ALD on an interfacial SiO_2 layer grown by chemical oxidation. Selected samples have received a PDA anneal either in N_2 at 600°C or in O_2 at 500°C. N-type poly-Si was used as gate electrode activated at 1000°C for 10 sec. All MOS devices have received a final forming gas anneal either at 400 or 420°C. Irrespective of the processing condition comparable electrical results were obtained for FETs and MOS capacitors. Typical results are presented and discussed below.

6.2.1 Hysteresis sensed by conventional techniques

The conventional hysteresis measurement discussed in Section 5.1.2 presents a fast screening method to study instabilities in high-ε gate dielectrics. In case of MOS capacitors the HF-CV technique was used to monitor the ΔV_{MG} shift, whereas for FET structures the I_D-V_G characteristic was used to monitor changes in the V_T.

In Fig. 6.20 typical results obtained from multiple I_D-V_G traces are shown for a variety of gate stacks with different interfacial SiO_2 layer thicknesses. As can be seen, the shift in V_T strongly depends on the interfacial SiO_2 thickness and the thickness of the HfO_2 layer. When a thin interfacial SiO_2 layer is used a shift in the V_T can already be observed at operation condition, whereas for thick interfacial layers the instability is retarded. In this case, the instability starts as soon as carriers are being injected by F-N tunneling through the thick interfacial layer. To illustrate the severity of the charge trapping in the HfO_2 layer at low bias conditions the SiO_2 reference is shown for comparison. In SiO_2 there is no measurable V_T shift up to $\sim 10MV/cm$ oxide field, which is estimated using $(V_G - V_T)/EOT$. For oxide fields above $10MV/cm$ the instability in SiO_2 is attributed to charge trapping in newly formed defect sites.

A qualitative assessment of the stability of the trapped negative charge for a n-channel MOSFET with a SiO_2 / HfO_2 gate stack is given in Fig. 6.21. The largest voltage shift compared to the fresh I_D-V_G characteristic is measured when the I_D-V_G characteristic is directly measured on the down trace from the maximum positive gate bias. When the I_D-V_G characteristic is sensed on the subsequent up trace starting at 0V gate bias a significant fraction of the trapped charge gets detrapped which results in a partial recovery of the V_T. To obtain complete charge recovery a sufficient negative gate bias has to be applied to the gate stack. The bias condition for complete discharging depends on the interfacial SiO_2 thickness and the discharging time.

The build up of negative charge in the SiO_2 / HfO_2 dual layer gate stack is also observed in the current - time trace during CVS. In Fig. 6.22 typical examples for current - time traces are shown. A recovery in the gate current is also evident when interrupting the stress for I_D-V_G sensing. This behavior is very similar to the one observed for *annealed* Al_2O_3 layers in Section 6.1. The current decay, however, cannot easily be converted into an absolute voltage shift when direct tunneling becomes the dominant leakage mechanism. Therefore, conventional measurement

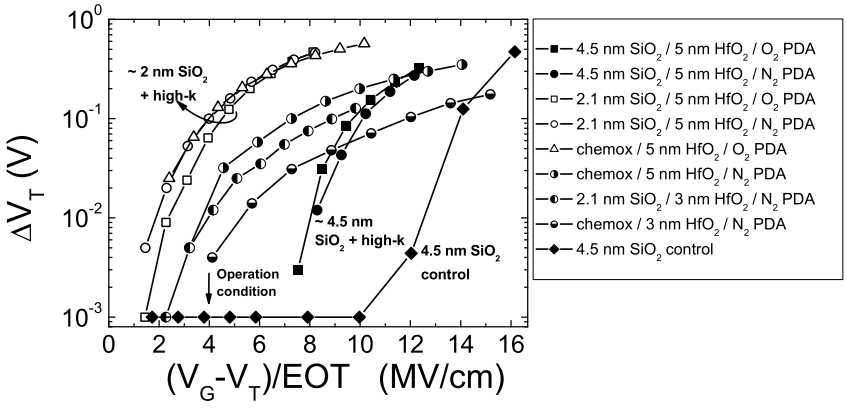

Figure 6.20: Stress induced V_T shifts versus maximum Si field for various HfO_2 gate stacks with different interfacial SiO_2 thickness extracted from multiple I_D-V_G traces as shown in Fig. 5.4 of Section 5.1.2. A 4.5 nm SiO_2 reference is shown for comparison.

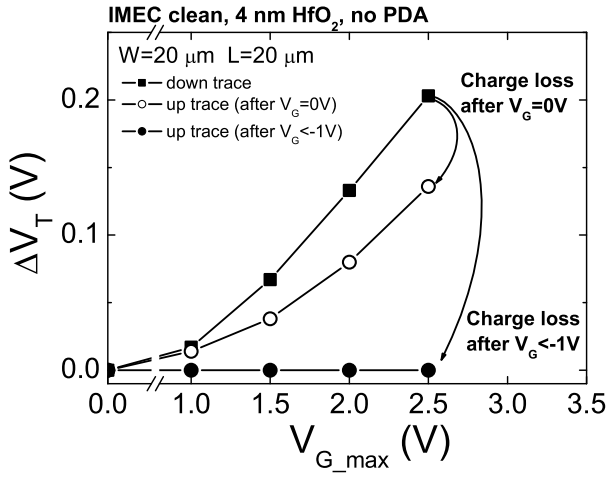

Figure 6.21: ΔV_T versus maximum gate bias V_{G_max} measured on the reverse trace and the subsequent up trace. Substantial charge loss after applying 0V to the gate and complete V_T recovery after applying a sufficient negative bias.

techniques were applied to qualitatively assess the charge trapping in SiO_2 / HfO_2 dual layer gate stack. Electron injection from the Si substrate is the main cause for the observed charging instability in n-channel MOSFETs with SiO_2 / HfO_2 gate dielectrics, similar to what was observed for Al_2O_3 layers. To correctly quantify the magnitude of the charge trapping novel time resolved measurement techniques in the ms to μs time range are required, as will be discussed next in Section 6.2.2.

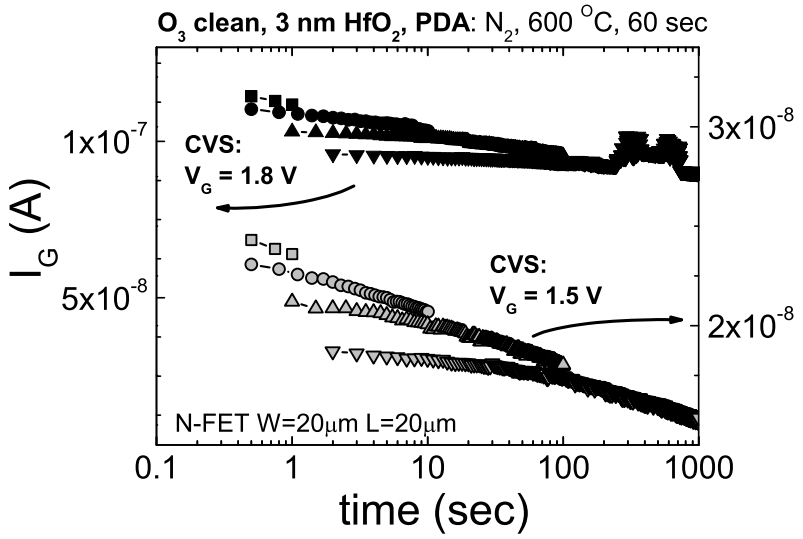

Figure 6.22: Gate current versus time during CVS stress measured on a n-channel MOSFET with a 3 nm HfO_2 layer. Current decay indicates build up of negative charge during stress.

6.2.2 Hysteresis sensed with time resolved techniques

Charge trapping in high-ε gate dielectrics is commonly studied using conventional measurement techniques like "stress and sense" or hysteresis measurements as described in Section 5.1. However, in the presence of transient charging effects the magnitude of the instability will not be correctly assessed by these techniques due to the inherent time delay between stressing and sensing or the slow ramp rates, which cause dynamic trap up. To overcome these issues the alternative fast measurement techniques, discussed in Section 5.2, are applied in this section.

In the pulsed I_D-V_G technique the measurement time is reduced by more than 3 decades compared to the conventional quasi DC ramp which usually takes ∼1s. When applying ramp rates of more than 10 kV/s the charge loss during the measurement is reduced substantially. In Fig. 6.23 a typical example for a pulsed I_D-V_G characteristic is shown for rise / fall times of 100 μs. The charging time was varied in this case from 100 μs to 500 ms. As can be seen, with increasing charging time the V_T continuously increases due to the build up of negative charge in the SiO_2 / HfO_2 dual layer stack. After stressing the device at $V_G = 2V$ for 500ms a change in V_T of $\sim 500mV$ is observed. Furthermore, when comparing the down trace for different stressing times no significant stretch-out is observed, as extracted from the falling edge of the trapezoidal pulse. From the traces shown in Fig. 6.23 it is evident that charge trapping in SiO_2 / HfO_2 dual layer stacks already happens in the μs time range.

To further illustrate the transient charging effects and its impact on the measured V_T shift, pulsed I_D-V_G measurements were carried out using a constant charging time of 100 ms and fall times ranging from 100 μs to 100 ms. As can be seen in Fig. 6.24, a more positive V_T is measured when the fall time is reduced towards the 100 μs range. This effect is attributed to detrapping during the down trace which is more pronounced for slower ramp rates. The magnitude of the measured V_T shift

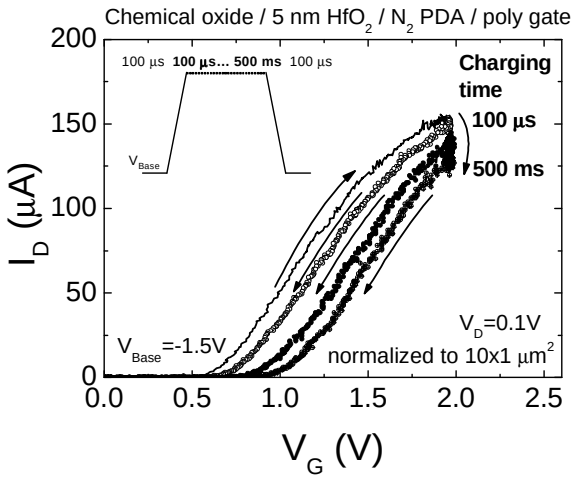

Figure 6.23: Pulsed I_D-V_G characteristics measured using charging times ranging from 100 μs to 500 ms and a fixed rise / fall time of 100 μs. Increasing V_T shift is observed with increasing charging time.

can increase by a factor of 2 and more when fast sensing techniques are applied.

The pulsed I_D-V_G characteristics shown previously were obtained by applying a single pulse to the gate of the n-channel MOSFET using a negative base level. However, during operation of such devices the gate bias will be either 0V or V_{DD}. To discuss the effect of the low level of the voltage pulse on the hysteresis measured by the pulsed I_D-V_G technique a comparison between single and repetitive pulses was made where either 0V or -2V was used as low level. As can be seen in Fig. 6.25 when a sufficiently negative bias is applied to the gate of the device complete discharging occurs even in the μs time range. Complete discharging results in an identical hysteresis trace for both single and repetitive pulses (compare square symbols in Fig. 6.25).

On the other hand, if a low level of the gate bias of 0V is used the hysteresis traces become quite different. For a single pulse applied to a fresh device, a similar trace is obtained compared to the ones discussed above. When repetitive pulses are used, however, complete discharging is not reached anymore and therefore charging starts to dominate over discharging. This effect results in a net positive voltage shift and the continuous closing of the hysteresis trace.

The pulsed I_D-V_G technique clearly demonstrates the presence of strong transient charging effects in the μs time range. For a reliable assessment of charge trapping these effects need to be taken into account.

6.2.3 Capacitance transients

Capacitance transients can also be used to study time dependent charging effects of the dielectric. In this measurement the LCR meter is simply used in a sampling mode using a constant DC bias. The time resolution of the measurement is basically limited by the measurement speed of the meter. For short integration time (low resolution) $\sim 100\,ms$ elapse between successive readings whereas for long integration time (high resolution) a measurement time of $\sim$1s is needed. Considering the time resolution of the instrument, transient effects slower than $\sim 100\,ms$ can be monitored.

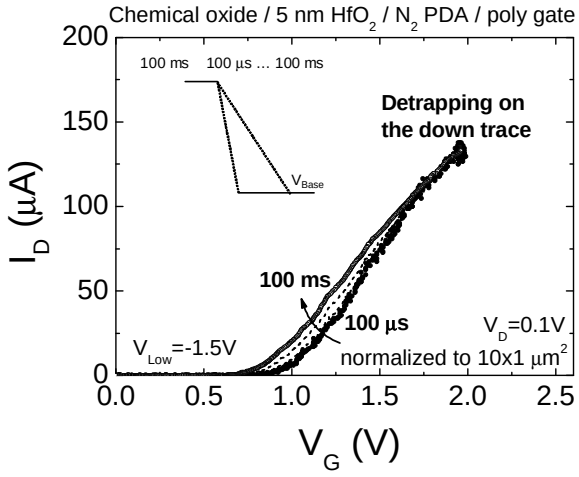

Figure 6.24: Down trace of the pulsed I_D-V_G characteristic as function of ramp rate using a fixed charging time. Detrapping during the down ramp clearly evident from the stretch-out of the I_D-V_G characteristic.

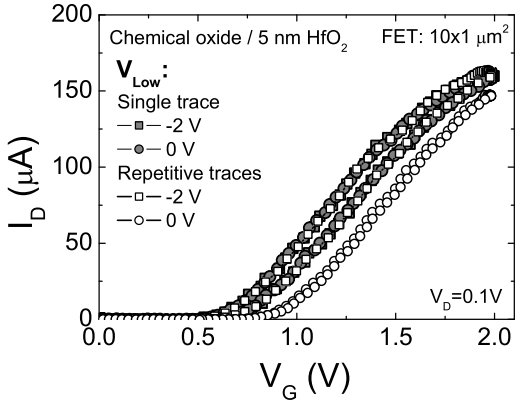

Figure 6.25: Single versus repetitive pulsed I_D-V_G traces using setup shown in Section 5.2.3. Complete discharging is only obtained when a sufficient negative gate bias is applied.

In Fig. 6.26 typical capacitance transients are shown for SiO_2 / HfO_2 gate stacks either on n- or p-type Si substrates with p- and n-type poly Si electrodes, respectively. As can be seen, when measuring the accumulation capacitance of NMOS and PMOS capacitors as function of time a significant decay in the normalized capacitance is observed in both cases. The normalized capacitance was derived from the data by calculating $C_{OX}(t)/C_{OX}(0)$. The capacitance decay is again attributed to the build up of negative charge for positive gate bias, whereas for negative gate polarity the decay is either caused by detrapping of negative charge or by trapping of positive charge. The transient effect is completely negligible in the SiO_2 reference sample, indicating a stable dielectric.

For the PMOS device electron injection from the n-type Si substrate occurs at

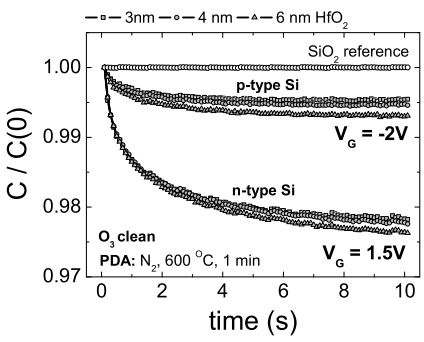

Figure 6.26: Normalized C-t characteristic of SiO_2 / HfO_2 gate stacks on n- and p-type Si substrate with p- and n-type poly Si electrode, respectively. SiO_2 reference is shown for comparison.

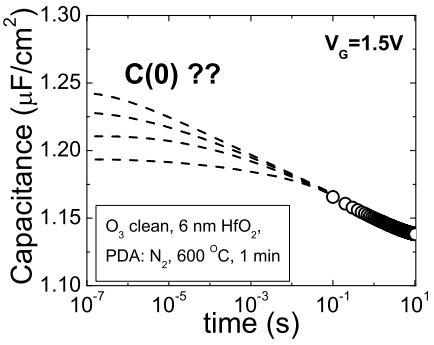

Figure 6.27: Capacitance time trace measured of a 6 nm HfO_2 layer on n-type Si substrate with p-type poly Si gate plotted on a logarithmic time scale. Conventional measurement procedures cannot resolve the "time zero" capacitance.

positive gate bias. Trapping of negative charge leads to a positive V_{FB} shift and therefore, according to the C-V swing the accumulation capacitance decreases. The capacitance change can be converted directly into a voltage shift using a simulated C-V characteristic. In case of the NMOS device the decay in the accumulation capacitance correlates with a negative V_{FB} shift. The negative shift can be caused either by hole trapping or detrapping of negative charge. Hole trapping is usually reported in combination with hot carrier injection which is not likely to be the case at the bias condition considered for the capacitance transient. Therefore, detrapping of negative charge is more likely the cause for the capacitance transient observed for the NMOS device.

Furthermore, the transients depend strongly on the polarity, but only weakly on the thickness of the HfO_2 layer. Such dependence further emphasizes the high efficiency of electron trapping under positive gate bias.

When plotting the capacitance transient on a logarithmic time scale, as shown in Fig. 6.27, it is evident that the "time zero" capacitance cannot easily be obtained by a conventional HF C-V measurement due to the presence of fast transient charging effects. To capture the initial phase of the charging phenomena time resolved techniques in the μs to ms time scale are required.

6.2.4 Drain current degradation

The presence of fast transient charging effects, as discussed in Section 6.2.3, are also observed in the drain current transients. When using a semiconductor parameter analyzer in the conventional sampling mode, transient effects, typically in the time range above 10 to 50 ms, can be recorded. The minimum time resolution is given by the measurement speed of the analyzer, which depends on the current resolution. The time scale towards shorter measurement times can be extended by using the pulsed technique as described in Section 5.2.3. In this case, instead of applying a trapezoidal pulse to the gate a square pulse is used and the drain current transient is extracted from the trace of the voltage drop at the load resistor. The same equation applies as given in Section 5.2.3.

Figs 6.28 and 6.29, show that by combining the conventional sampling technique with the pulsed measurement the onset of charge trapping can be resolved. The

data shown in the figures illustrate the importance of the interfacial SiO_2 layer thickness for the trapping behavior of the SiO_2 / HfO_2 dual layer stack. A thicker interfacial SiO_2 layer results in a delayed onset of charge trapping for a given gate bias, especially in the low voltage regime. This can be seen when the traces for $V_G = 1.25V$ in Fig. 6.28 and 6.29 are compared. However, when the gate bias is increased and thereby enhanced carrier injection occurs the delayed onset becomes less evident. At high bias condition charge trapping sets in already after $\sim 10\mu$s.

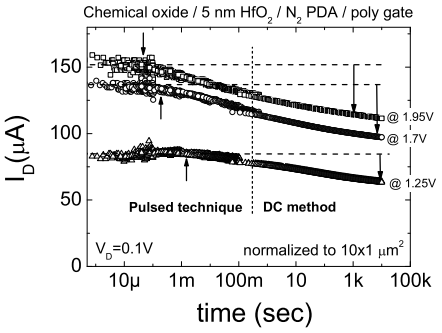

Figure 6.28: Drain current as a function of time for a 5 nm HfO_2 layer grown on a chemical oxide starting surface using stress bias as indicated in the figure. Combining pulsed measurements with the conventional techniques allows to monitor the drain current degradation over more than 10 decades in time.

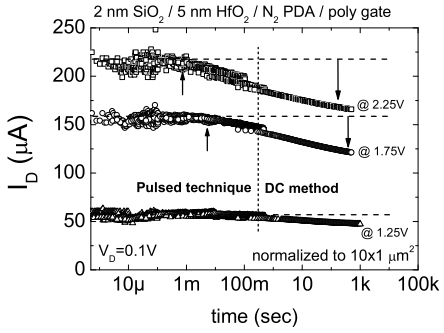

Figure 6.29: Drain current as a function of time for a 5 nm HfO_2 layer grown on 2 nm SiO_2 interface using stress bias as indicated in the figure. Pulsed measurements are combined with the conventional techniques to monitor the drain current degradation.

The drain current degradation can be converted into a V_T shift using the pulsed I_D-V_G characteristic. In this case the current degradation is solely attributed to a change in the threshold voltage. Second order effects caused by enhanced scattering due to trapped charge are neglected here. As shown in Fig. 6.30, the current degradation can simply be converted into a voltage shift and then be compared to the conventional "stress and sense" procedure. In Fig. 6.31 the comparison between "stress and sense" and "sense at stress" is given. As can be seen, $\sim$ 2x larger instability is measured when using the drain current degradation to monitor the charge build up in the dielectric. This enhancement is attributed to the fast transient charging and discharging effects which are not captured by the conventional procedures.

Not only the magnitude of the instability is important for a correct prediction of the instability, crucial is also the time dependence. A lifetime prediction for charge trapping definitely needs to consider the fast transient charging effects.

At this point it can only be speculated what the impact of further dielectric scaling or changing to metal gate electrodes on the charge trapping and especially on the voltage and time dependence will be. In general, reducing the thickness of the high-ε material will also reduce the number of defect sites available to be charged by electrons. Furthermore the balance between trapping and detrapping of negative charge may shift towards detrapping and therefore effectively less charge remains trapped in the dielectric. On the other hand, scaling will enhance charge injection which might be detrimental from the trapping point of view. However, in all cases scaling will reduce the impact of a single carrier on the V_{FB} or V_T instability due to increasing oxide capacitance. It remains an open question, whether scaling by itself will sufficiently reduce the instabilities and therefore the demand to reduce the

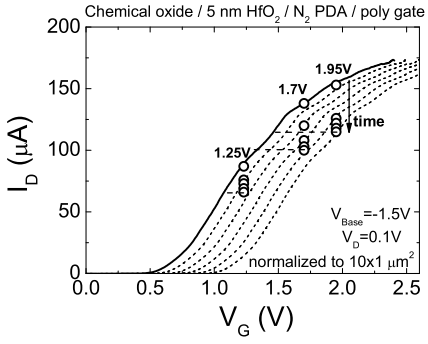

Figure 6.30: Drain current degradation converted into V_T shift using pulsed I_D-V_G characteristic. Current degradation due to 2^{nd} order effects are neglected.

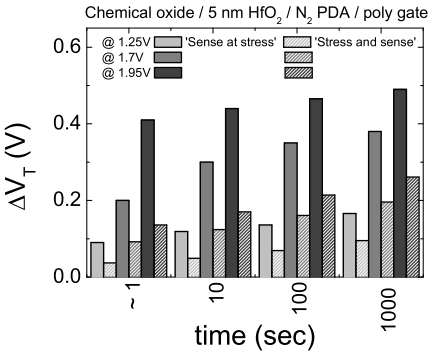

Figure 6.31: Comparison of ΔV_T measured using conventional "stress and sense" and "sense at stress". The conventional technique significantly underestimates the the effect due to transient charging effects.

defect density in the high-ε material may be relaxed at least for high performance application.

6.2.5 Charge trapping monitored by the C-P technique

Charge pumping has been introduced as a technique to study the properties of the interface between the Si substrate and gate dielectric in conventional MOSFETs. The application of the charge pumping technique, however, to MOSFET with high-ε dielectrics has not been described in detail so far. In this section, an attempt is made to explain the charge pumping characteristics when charging and discharging of bulk defect sites in combination with interface traps contribute to the measured substrate current. Geometric effects will also be discussed.

The basic principle of charge pumping has been discussed already in Sections 2.2.3 and 5.2.5. To separate the contributions related to interface states from that associated with bulk defects the conventional basel level charge pumping with a small amplitude was applied. A comparison of the interface state density measured on a SiO_2 / HfO_2 dual layer stack and the SiO_2 reference is given in Fig. 6.32. When a voltage pulse with a small amplitude ($V_{Amp} = 1.2V$) is used to measure the charge pumping characteristic, the interface state densities (N_{CP}) of the SiO_2 / HfO_2 dual layer stack are similar to that of the oxide control. Furthermore, the charge pumping current scales well with frequency indicating no significant contribution from slow charging and discharging effects under these bias conditions. The data indicate that the interface properties of the dual layer stack are quite similar to single layer SiO_2 and that the conventional passivation process is sufficient for the short channel device considered in this experiment.

When an amplitude sweep instead of the conventional base level sweep is applied to the SiO_2 / HfO_2 dual layer stack the contribution from charging and discharging of bulk defect states becomes clearly evident, as shown in Fig. 6.33. The normalized charge per cycle strongly increases with increasing peak level and charging time. This behavior is a strong indication that the defect sites which contribute to the recombination current in the substrate are distributed both in energy and space. For the oxide reference basically no voltage dependence is observed.

The influence of the interfacial layer on the charging and discharging of bulk

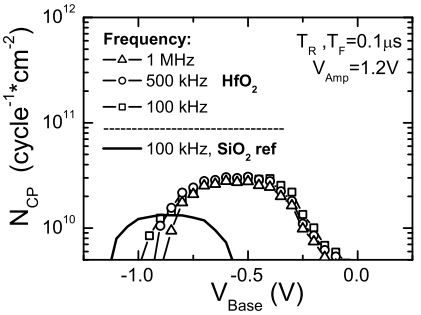

Figure 6.32: Charge per cycle measured by C-P on SiO_2 / HfO_2 dual layer stack using the parameters given in the figure. Low interface state density indicates SiO_2 like behavior with well passivated interface. SiO_2 reference is given for comparison.

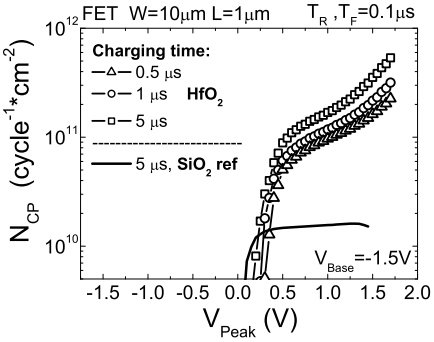

Figure 6.33: Charge per cycle measured by C-P on SiO_2 / HfO_2 dual layer stack using the amplitude sweep and the parameters given in the figure. A strong increase in N_{CP} due to charging / discharging of bulk defects in the HfO_2 layer is observed.

defect sites in a charge pumping measurement is illustrated in Figs. 6.34 and 6.35. When comparing the two figures, it is quite evident, that the thickness of the interfacial SiO_2 layer plays an important role in charging and discharging of bulk defects in the HfO_2 layer. When the HfO_2 layer is deposited on a thin chemical oxide charge trapping takes place already at low voltages and short charging times. Whereas a thicker interfacial SiO_2 layer substantially suppresses the fast component in the charge trapping, as shown in Fig. 6.35. Another interesting feature of the charge trapping is given by the comparison between the recombination current measured at the substrate and the current at source / drain (compare open and solid symbols in Fig. 6.34 and 6.35). At low injected charge densities the source / drain current is equal to the substrate current suggesting that all the carriers injected into the gate stack are being emitted back to the Si substrate when the bias is reversed. When the amplitude of the pulse is increased the current traces start to deviate indicating either charge detrapping to the gate electrode or the contribution of a tunneling current form the inversion layer to the gate electrode.

The influence of the channel length on the charging / discharging of bulk HfO_2 defects was studied in more detail. To separate the contribution from interface states and bulk defect states to the substrate current, the conventional base level sweep and amplitude sweep were performed on devices ranging in channel length from $<1\,\mu$m to $100\,\mu$m. The rise / fall time was varied from 100 ns to $100\,\mu$s in order to vary the thermal emission of trapped interface charge and to observe the contribution from the geometrical component to the substrate current. As shown in Fig. 6.36, the charge per cycle is reduced in a conventional charge pumping measurement when lower rise / fall times are used which is consistent with thermal emission of trapped interface charge. Therefore, the charge per cycle can be normalized to a nominal rise / fall time using Eq. 2.44 in Section 2.2.3 where a uniform interface trap distribution is assumed and the capture cross sections for electrons and holes are taken from the literature [14]. The bump in the conventional charge pumping characteristic in Fig. 6.36 is caused by charging / discharging of bulk defect sites. To minimize this effect a small pulse amplitude has to be used as discussed previously.

From data, as shown in Fig. 6.36, the substrate current was converted into a interface trap density with a nominal rise / fall time of 100ns using Eq. 2.44. Fig.

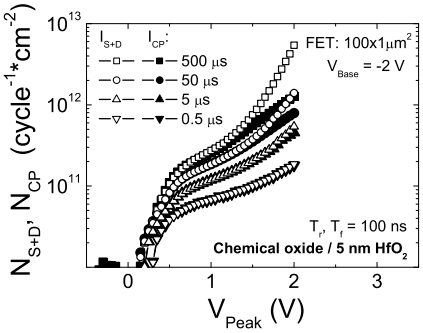

Figure 6.34: Comparison of charge per cycle measured at the Si substrate and Source / Drain on a 5 nm HfO_2 layer grown on a chemical oxide starting surface. At low charge densities $I_{CP} = I_{S+D}$ whereas at high densities $I_{CP} < I_{S+D}$.

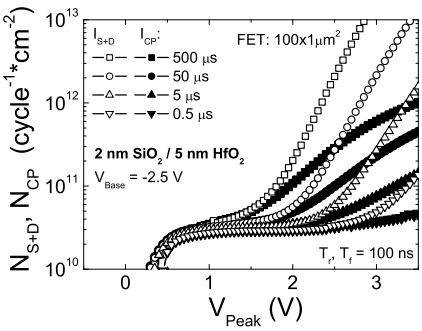

Figure 6.35: Comparison of charge per cycle measured at the Si substrate and Source / Drain on a 5 nm HfO_2 layer grown on 2 nm SiO_2 interface. Same general behavior as shown in Fig. 6.34, however, dependence on voltage and charging time differs significantly.

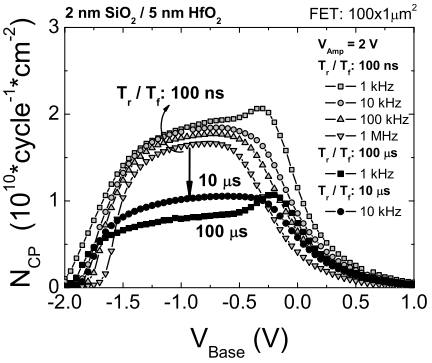

Figure 6.36: Charge per cycle as function of frequency and rise / fall time using conventional base level sweep. The reduction in I_{CP} is attributed to thermal emission of trapped interface charge as explained by the charge pumping theory discussed in Section 2.2.3.

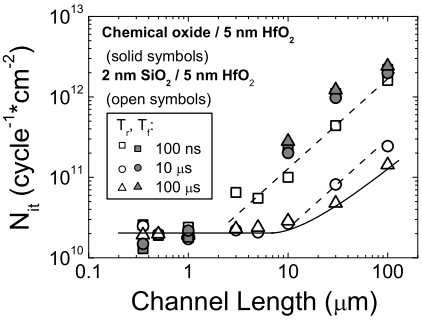

Figure 6.37: Channel length dependence of N_{it} measured using the conventional base level and conditions as indicated in the figure. The interface trap density was normalized to $t_r = t_f = 100$ ns correcting for thermal carrier emission and assuming a uniform trap distribution as given by Eq. 2.44.

6.37 illustrates that for short channel devices ($< 10\,\mu$m) low interface state densities are extracted independent of the interfacial layer thickness. When the channel length increases, however, differences in the extracted interface state densities are obtained. In case of the 2 nm SiO_2 the apparent increase in interface state density for long channel devices ($> 10\,\mu$m) and for short rise / fall times is due to recombination of inversion carriers [17]. No dependence on the rise / fall time is observed for the HfO_2 sample with the thin chemically grown SiO_2 layer, indicating enhanced interface trap densities for long channel ($> 10\,\mu$m) devices [57].

To investigate the channel length dependence of the bulk trapping in SiO_2 / HfO_2 dual layer stacks the amplitude sweep was used with a frequency of 1 kHz and 2.5 kHz. A long rise / fall time was chosen to minimize the geometrical component. In Fig. 6.38 the charge pumping characteristic of the amplitude sweep is shown for devices with different channel length. For short channel devices a strongly reduced effect is measured. When the normalized charge per cycle, $N_{CP} = (N_{CP} - N_{it})/(N_{CP} - N_{it})_{max}$, after correction for interface state contributions, is plotted versus channel length, as shown in Fig. 6.39, a clear roll off is observed independent of the interfacial SiO_2 thickness. The roll off behavior of the normalized charge per cycle is caused by charge loss to the source / drain junctions. When the trapped electrons are emitted from the HfO_2 layer back into the p-type Si substrate of the n-channel MOSFET they can diffuse into the substrate prior to recombination with majority carriers. The diffusion of electrons back to the source / drain junctions becomes more effective in devices with shorter channel length, which leads to a reduction of the recombination current measured at the Si substrate [58]. For a correct assessment of the charge trapping using the charge pumping technique, this effect has to be included accordingly. For monitoring improvements with respect to process development, measurements on devices with identical geometry are necessary.

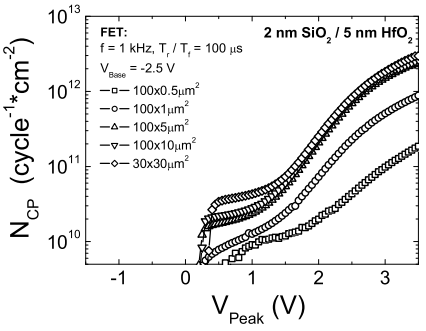

Figure 6.38: Charge pumping characteristic versus channel length using amplitude level charge pumping. The current measured at the Si substrate strongly decreases with decreasing channel length. Long rise / fall times were used to minimize the geometrical component.

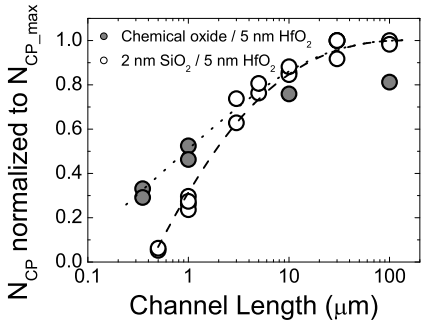

Figure 6.39: Normalized N_{CP} versus channel length obtained from data as shown in Fig. 6.38 corrected for N_{it} as shown in Fig. 6.37. The decrease in normalized N_{CP} with channel length is due to enhanced charge collection by the junctions.

Finally, the finding of enhanced interface state density in large area devices was directly confirmed by the pulsed C-V method as shown in Fig. 6.40. The distortion in the C-V characteristic is caused by charging and discharging of the interface states when sweeping the device from accumulation to inversion and vice versa. From the additional displacement current a interface state density of $\sim 5 \cdot 10^{11} cm^{-2}$ is extracted which is in agreement with the numbers given in Fig. 6.37 for the 2 nm SiO_2 interface.

In summary, charge pumping was introduced to study interface and bulk defects in SiO_2 / HfO_2 dual layer gate dielectrics. The conventional base level sweep with a rather small amplitude was used to sense interface states, whereas the amplitude sweep was applied to investigate the bulk defects in the HfO_2 layer. For a correct assessment of the bulk defect density the geometrical effects need to be considered carefully.

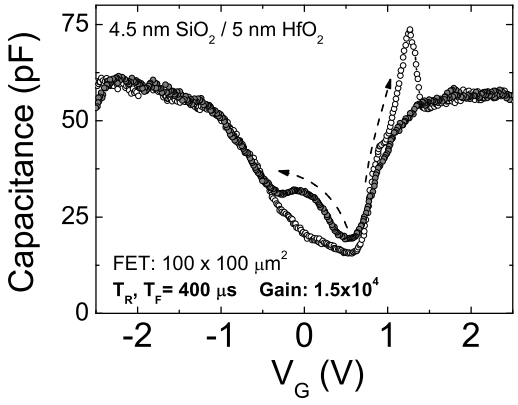

Figure 6.40: Pulsed C-V characteristic measured on a large area FET. The presence of interface states is directly confirmed by the C-V measurement.

6.2.6 Defect band model

The data presented in Section 6.2.1 to 6.2.8, can be summarized as follows:

- Charge trapping in the bulk of the HfO_2 layer will contribute to the V_T instability typically observed in scaled stacks. Even aggressively scaled stacks will contain 3 to 4 nm thick HfO_2 layers. Based on the presented data, it is expected that much of the bulk trapping properties will contribute to the instability of scaled layers. From the volume reduction, a reduction of charge trapping by less than a factor of 5 can be expected. Additional benefits may arise from reduced charge injection at low voltages near operational conditions and from enhanced detrapping because of the thickness reduction. The contacts are closer to the centroid of the trapped charge, which may enhance detrapping. The possible benefit of these effects need to be evaluated in scaled stacks. A remaining concern is certainly the fact that the charging is of transient nature, which causes V_T instabilities on very short time scales. Improving the bulk properties of high-ε materials will be of crucial importance for succesful integration. Possible improvements may come from optimized post deposition treatments or from alternative deposition conditions / methods.

- The charging behavior observed in the thick stacks is qualitatively consistent with the trapping behavior inferred from V_T instabilities in scaled stacks. Based on measurements in scaled stacks, a trapping / detrapping model is proposed which can account for many of the qualitative aspects of the the trapping / detrapping behavior in SiO_2 / HfO_2 gate stacks. As illustrated in Fig. 6.41, a defect band is proposed in the HfO_2 layer, which supplies the electronic states for the observed electron trapping. Especially for thin interfacial layers, this model has some attractive features. First, the band is not occupied or only lightly occupied at flatband conditions, resulting in small initial V_T and V_{FB} shifts. The position of the band with respect to the Si conduction band is predicted to vary strongly with gate bias:

$$d\mathcal{E}/q = V_G \cdot t_{SiO_2}/EOT \tag{6.8}$$

This is enforced by the differences in dielectric constant of the two constituents of the layer. For positive gate bias, the band moves below the Fermi level in the Si substrate and the band can be filled by tunneling into the shaded area (Fig. 6.41c) and possibly by transport through the defect band. For negative gate bias, the states in the shaded area in Fig. 6.41a will be emptied by tunneling. Transport from the bulk into the shaded area may also be present for negative bias. This feature explains the observed polarity dependence. Some issues remain with respect to charge trapping under negative gate bias. The asymmetry in the gate current density may be of importance to fully understand the strong asymmetry with polarity, including the charge reversibility observed here.

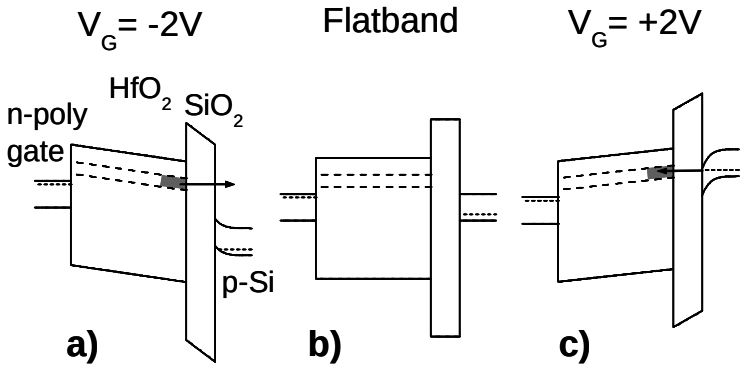

Figure 6.41: Schematic energy band diagram of a SiO_2 / HfO_2 / poly-Si gate stack containing a defect band in the HfO_2 layer at flatband condition b). For negative a) and for positive gate bias c), the defects located near the SiO_2 interfacial layer move rapidly with respect to the Fermi-level in the Si substrate.

- The exact origin of the defect band remains speculative. Oxygen vacancies have been proposed to be the cause in other oxides [59]. It is also possible, that water related defects such as OH (OH^-) could be the root cause. Additionally, the films contain Cl impurities at the few percent level. These inputs provide some guidance for possible improvements. The density of oxygen vacancies might be reduced by oxygen annealing. However, rapid interfacial oxide growth strongly limits the acceptable thermal budget for this approach. Water related species might be reduced by high temperature anneals in an inert ambient. Such anneals could be followed by low temperature oxygen soak anneals. Finally, the use of other deposition methods with other precursors should allow the investigation of the impact of Cl via a comparative study.

6.2.7 The injector concept

Strong transient charging effects are observed in dual layer dielectrics with a thin interfacial SiO_2 as described in the previous sections. These transient effects are mainly caused by tunneling of carriers through the interfacial layer into or from the bulk of the high-ε material. The tunneling process can efficiently be suppressed when a thick SiO_2 layer is used as interface. The injection of carriers (electrons) in that case is controlled by F-N tunneling through the SiO_2 layer.

The schematic drawing of the injector concept is given in Fig. 6.43. Carrier injection through the thick interfacial SiO_2 is controlled by F-N tunneling which is negligible at oxide fields below 6MV/cm. In case of thin interfacial layers tunneling currents are already measured at a low oxide voltage, as illustrated in Fig. 6.42. To separate the impact of carrier injection from the field dependence of the observed instabilities and, furthermore, to eliminate the strong transient charging effects due to tunneling through the thin interfacial layer, the injector concept has been proven to be a very useful method.

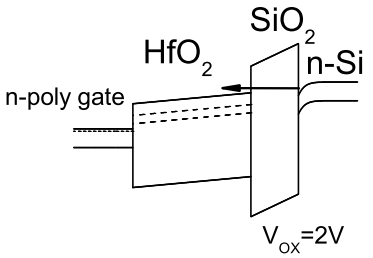

Figure 6.42: Schematic energy band diagram of a 2 nm SiO_2 / 5 nm HfO_2 gate stack for positive gate bias.

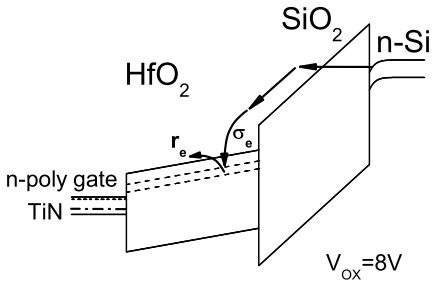

Figure 6.43: Schematic energy band diagram of a 9 nm SiO_2 / 12 nm HfO_2 gate stack for positive gate bias. Charge injection is controlled by F-N tunneling.

Injector stacks were fabricated on n-type Si substrate with either poly-Si or metal (TiN) electrodes. An interfacial SiO_2 thickness of either 5 or 9 nm was chosen followed by the HfO_2 deposition of 6 or 12 nm for the poly or metal stack, respectively. In Figs. 6.44 and 6.45 the current voltage characteristic of the injector stacks are shown and compared to the SiO_2 reference. As can be seen, there is no measurable gate current in the low voltage regime due to the presence of the thick interfacial SiO_2 layer. When the gate voltage is increased the F-N current starts and well behaved I-V characteristics are measured for the control samples. The samples with SiO_2 / HfO_2 dual layer dielectric also show well behaved F-N tunneling characteristics at low current densities, but for high current density a strong build up of negative charge is observed. This effect becomes more clear, when the gate current is plotted versus the injection field as shown in Figs. 6.46 and 6.47. In all cases studied, the high-ε material can easily trap charges in excess of $N_{OX} > 10^{12} cm^{-2}$.

To discuss the correlation between charge injection and trapping the injector concept was used and combined with a constant current and constant voltage stress procedure. As can be seen in Fig. 6.48, charge trapping in SiO_2 / HfO_2 gate stacks is very different from furnace SiO_2. In SiO_2, during constant current injection the gate voltage is constant as soon as the accumulation layer is formed and F-N injection occurs. During stress, SiO_2 exhibits charge trapping only at high fluence ($N_{inj} \approx 10^{16}/cm^2$). In contrast, in SiO_2 / HfO_2 stacks, very fast charge trapping takes place already at low fluence, $N_{inj} < 10^{14}/cm^2$. This behavior is observed independent of the PDA condition. From the data in Fig. 6.48, a defect density, $N_{tr} > 10^{12}/cm^2$, is estimated regardless of the PDA conditions. Similar results are obtained for dual layer stacks with poly Si electrodes [60].

To better quantify the trapping transients, the conventional stress and sense procedure was applied to the injector stacks whereas the time-resolved techniques

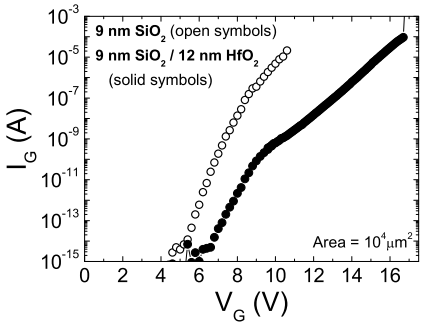

Figure 6.44: Gate I-V comparison between the 9 nm SiO_2 reference and a 9 nm SiO_2 / 12 nm HfO_2 dual layer gate stack with poly-Si electrode.

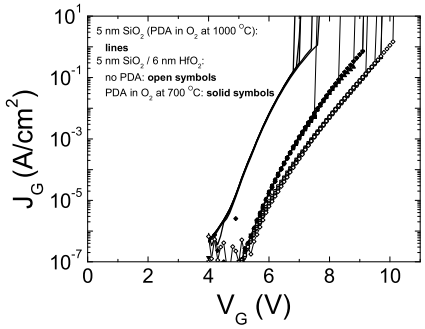

Figure 6.45: Gate I-V comparison between the 5 nm SiO_2 reference and a 5 nm SiO_2 / 6 nm HfO_2 dual layer gate stack with TiN gate electrode.

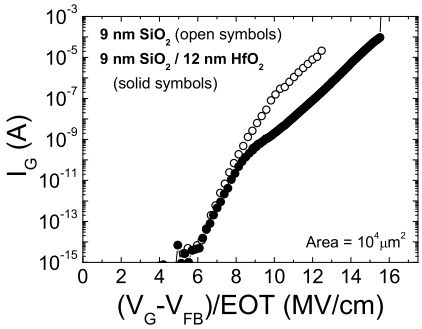

Figure 6.46: Gate current versus injection field for the 9 nm SiO_2 reference and the 9 nm SiO_2 / 12 nm HfO_2 dual layer gate stack with poly-Si electrode. No deviation in gate current is found for low injection conditions between the different stacks, whereas for high injection conditions the difference becomes evident.

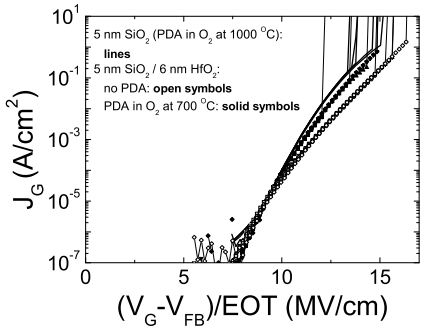

Figure 6.47: Gate current versus injection field for the 5 nm SiO_2 reference and the 5 nm SiO_2 / 6 nm HfO_2 dual layer gate stack with TiN gate electrode. Same general behavior as for the poly gate devices shown in Fig. 6.46. Only small differences in electron trapping are observed depending on the post deposition treatment.

were used in addition for scaled gate stacks. In the following, the charge trapping behavior of thin and thick stacks will be compared.

For a given amount of injected charge, N_{inj}, the apparent trapping probability,

$$\overline{P(N_{inj})} = \overline{x}P(N_{inj}) \tag{6.9}$$

is proportional to the trapping probability, P, and the charge centroid, $\overline{x}$, as measured from the gate electrode. The apparent trapping probability, $\overline{P}$, was determined using:

$$\overline{P(N_{inj})} = \Delta N_{tr}/\Delta N_{inj} \tag{6.10}$$

$$\Delta N_{tr}(N_{inj}) = C_{OX}\Delta V/q \tag{6.11}$$

$$\Delta N_{inj} = \langle J_G \rangle \Delta t/q \tag{6.12}$$

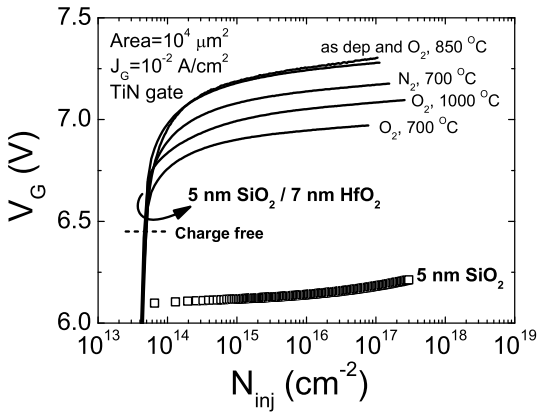

Figure 6.48: Gate voltage versus fluence, N_{inj}, during a constant current stress ($J_G = 10mA/cm^2$, substrate injection) for 5 nm SiO_2 (open square) and for dual layer stacks with a 7 nm thick HfO_2 layer on top of a 5 nm SiO_2 layer. TiN was used as gate electrode. In contrast to the 5 nm SiO_2 control, fast charge trapping is observed in the HfO_2 containing stacks, irrespective of PDA condition.

$$N_{inj} \quad = \quad \sum \Delta N_{inj} \tag{6.13}$$

The quantity, $\Delta V = V(t+\Delta t) - V(t)$, is the flatband voltage shift, $\Delta V = \Delta V_{FB}$, or the threshold voltage shift, ΔV_T, measured in capacitors and FETs, respectively. In the latter case ΔV_T was calculated from the drain current change at stress condition, using $\Delta V_T = \Delta J_D / g_m$, where g_m is the transconductance [61].

The measured apparent trapping probabilities, $\overline{P}$, for two different gate stacks are shown in Figs. 6.49 (using stress and sense) and 6.50 (using sense at stress in FETs). Independent of the SiO_2 thickness, a large trapping probability, $\overline{P(N_{inj})} \approx 0.3$ is measured for $N_{inj} < 10^{12}/cm^2$, and $\overline{P}$ decays with $(N_{inj})^{-1}$ at larger fluence. The fact that the value of $\overline{P}$ appears to be saturated at 0.2 to 0.3 can be explained by the charge location, $\overline{x}$, according to Eq. 6.9. If we assume that the charge is located in the bulk of the HfO_2, a charge centroid of $\overline{x} \sim 0.3$ is estimated, suggesting that all the injected charge gets trapped initially, $\overline{P} \approx 1$, providing a simple explanation for the saturation of $\overline{P}$.

Furthermore, since $\overline{P}$ is independent of the gate voltage, the trapped charge can simply be calculated by integrating $\overline{P}$ over the injected charge. As expected, this approach seems valid, as illustrated by the examples in Figs. 6.51 and 6.53. In Fig. 6.51, the measured trapped charge (solid symbols), not corrected for the charge location, is compared to the charge predicted by integrating the apparent trapping probability (open symbols) from Fig. 6.50. At voltages higher than 1.4V, the initial charging transient cannot be captured because of measurement speed limitations. (The reference drain current (1^{st} point measured) to extract the voltage shifts is already degraded). These limitations can be overcome by using time resolved I_D-V_G technique as introduced in Section 5.2.3. During a single pulse measurement, the drain and gate voltages are recorded simultaneously using a digital scope. From the voltage traces, the I_D-V_G characteristics is extracted for the up and down trace. The pulse shape (width and maximum V_G) determines the amount of injected charge. In Fig. 6.52 we compare the transient up and down traces for FETs with a 3 nm

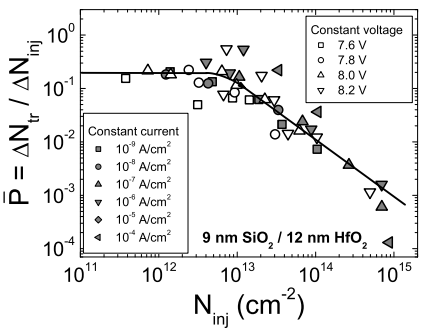

Figure 6.49: Apparent trapping probability $\overline{P}$ versus N_{inj} as measured with conventional stress and sense (C-V) method. Both, constant current stress and constant voltage stress were used for stressing, yielding similar results. At low N_{inj}, effective trapping probabilities, $\overline{P(N_{inj})} \approx 0.3$, are measured. (Note that minimal charge loss was observed in this stack with 9 nm SiO_2, justifying the use of a "stress and sense" technique in this case.)

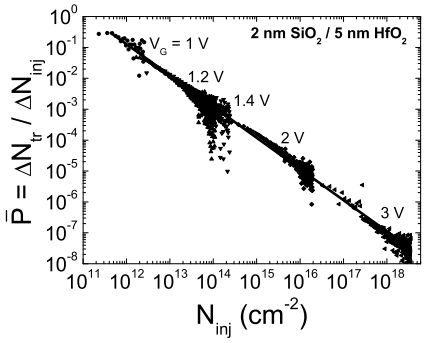

Figure 6.50: Apparent trapping probability, $\overline{P}$, as a function of electron fluence, N_{inj}, for several gate biases. At each bias condition the gate and drain current were measured simultaneously. Note that at low N_{inj}, $\overline{P(N_{inj})} \approx 0.3$.

HfO_2 layer on a 1 nm and 2 nm thick interfacial SiO_2 layer, respectively. At the selected stress voltage and charging time, no charge trapping is observed for the FET with the 2 nm thick interfacial layer. In contrast, the FET with the 1 nm thick interface layer shows substantial voltage shifts due to charge trapping. From Fig. 6.52, it is clear that for a given maximum V_G, the amount of trapped charge depends critically on the interfacial layer thickness.

This effect is quantified in Fig. 6.53, where the trapped charge measured by the pulse technique is compared for the two samples with different interfacial layer thickness as a function of the gate voltage using a constant charging time of $\Delta t = 100\mu s$. The solid lines show the predicted trapped charge based on the trapping probability measured in thick films, as shown in Fig. 6.50 and the estimated injected charge, $N_{inj} = \Delta t \cdot J_G$, where J_G is the measured gate current density. The trapping probability is integrated over the injected charge at each voltage condition. Charge trapping can reasonably well be predicted by this simple empirical approach. The interfacial SiO_2 controls the supply of electrons which can flow to the gate, and thus the amount of charge trapped at each voltage condition.

In summary, charge trapping in SiO_2 / HfO_2 gate dielectric was studied using thick stacks. The results suggest that efficient electron trapping occurs in the HfO_2 layer. The HfO_2 trapping probability does not seem to depend strongly on the PDA condition and the gate electrode used. A phenomenological trapping model was developed, which can qualitatively explain the charging phenomena including the effect of the interfacial layer thickness. The model is simply based on the electron fluence calculated from the gate current and the injection time.