17. CHARGE-BASED MEASUREMENT TECHNIQUES

17.1 Review/Background

In the previous lectures, systematic studies of NBTI and HCI mechanisms of trap formation were considered. The identification of the defect density that is formed is a major issue and there are various methods for extracting such information. The purpose of this lecture is to study three different methods for the measurement of the defect formation in MOSFET structures. These are the C-V method, the method of time-varying subthreshold slope, and the on-the-fly Idlin method.

17.2 Method No1: The C-V Method

Description of the C-V Measurement: Capacitance voltage measurements of MOS capacitors provide a wide range of different information about the MOS structure. Apart from the identification of quantities like the flatband voltage, the flatband capacitance, and the depletion width of a MOS structure, the nonideal properties of a MOS structure can be revealed, such as the fixed and mobile charge in the insulator and the existence of trapping interface states. A typical measurement result of C-V method is shown in Fig. 17.1.
In Fig. 17.2, one can observe the energy distribution of interfacial traps for oxidized silicon in directions <100> and <111>. The presence of these defects becomes observable through the stretch-out of the C-V characteristic which is explained in the following subsection.

**Explanation of C-V Stretch-out:** The C-V stretch-out is an accumulative effect of donor- and acceptor-like interfacial states. Both types of interfacial states contribute to a change in the threshold voltage \([i, ii]\). Referring to Fig. 17.3, a donor-like state is positively charged if it is empty and neutral when it is full. On the other hand, an empty
acceptor-like state is neutral, while it is negatively-charged upon being fully occupied. As a result, depending on the position of the Fermi level (indicated by the dashed line) which depends on the gate voltage, donor- and acceptor-like states show different behavior in different operation regime: accumulation, depletion, or inversion. The below expression gives the threshold voltage shift due to donor-like states.

\[
V_{th} = V_{th}^* - \frac{1}{C_{ox}} \int_{x_0}^{x_0} x a(V_g) Q_{ox}(x) \delta(x - x_0) \, dx \\
= V_{th}^* - \frac{a(V_g) Q_{ox}(x)}{C_{ox}}
\]

Here, \(a(V_g)\) accounts for the fraction of the interfacial states that contribute with nonzero charge to the threshold shift.

Fig. 17.3 (a) A donor-like state. (b) An acceptor-like state. (c) A combination of both the donor- and acceptor-like states. The Fermi level is shown by the dashed line.

The strong gate voltage dependence of this factor is hidden behind the perturbation in the C-V plot which is shown graphically in Fig. 17.4.
It should be noted that the number of Si/SiO$_2$ interface defects decreases upon forming gas annealing at high temperature (Fig. 17.5). This reduction is reflected in C-V curve through the decrease in the stretch-out amount (Fig. 17.6).

**Fig. 17.4** Effect of gate voltage modulation of threshold voltage for donor-like states.

**Fig. 17.5** (a) A schematic illustration of the effect of forming gas annealing on the passivation of the interface dangling bonds. (b) Decrease of the interface trap density ($D_{IT}$) as a result of annealing.
Representation of Defects by a Capacitance: A clearer evidence of the presence of defects comes through the application of stress. The stress affects both the C-V gate charge which is the sum of the interface charge ($Q_{IT}$) and the substrate charge ($Q_s$) which itself consists of the depletion region charge ($Q_B$), the electrons ($Q_n$), and the holes ($Q_p$). The gate voltage ($V_G$) consists of the voltage drop across the oxide ($V_{ox}$) and the voltage drop across silicon ($\psi_s$). Therefore, Eq.

$$C_T^{-1} = \left[ \frac{dQ_G}{dV_G} \right]^{-1} = - \frac{dV_{ox} + d\psi_s}{dQ_s + dQ_{IT}}$$

Fig. 17.6 C-V stretch-out decreases upon forming gas

Fig. 17.7 Effect of defects is
17.2 becomes:

\[ C_T^{-1} = -\frac{dV_{ox}}{dQ_G} - \frac{1}{dQ_s/d\psi_s + dQ_{IT}/d\psi_s} \]

\[ = \frac{1}{C_{ox}} + \frac{1}{dQ_B/d\psi_B + dQ_p/d\psi_p + dQ_n/d\psi_n + dQ_{IT}/d\psi_s} \]

\[ = \frac{1}{C_{ox}} + \frac{1}{C_B + C_p + C_n + C_{IT}(t)} \]

Derivative of the interface charge \((Q_{IT})\) with respect to surface potential \((\psi_s)\) gives the interface capacitance which is a function of time \((C_{IT}(t))\). Considering that \(C_n\) (capacitance due to the electrons, \(dQ_n/d\psi_s\)) and \(C_p\) (capacitance due to the holes, \(dQ_p/d\psi_s\)) are negligible, the last line in Eq. 17.3 translates to:

\[ C_T = \frac{C_{ox}(C_B + C_{IT}(t))}{C_{ox} + C_B + C_{IT}(t)} \]

Defect density using C-V analysis

Method A)

To extract the defect density from C-V measurement the following approach is used.

1) Given the device physical properties (oxide thickness, dimensions, etc.), the C-V curve for the ideal case (no defect, \(D_{IT}=0\)) is simulated as shown in Error! Reference source not found. (a).

2) Therefore, at each gate voltage the surface potential \((\psi_s)\) and the location of quasi-Fermi level \((F_N)\) are obtained for the ideal C-V (see Error! Reference source not found. (a)).

3) For each gate voltage, the occupancy of the interface defect density versus energy \((D_{IT}(E))\) is then calculated from the location of \(F_N\) which in turn gives the net charge and its corresponding shift in the gate voltage \((\Delta V_G)\). Then, shift the C-V curve by this gate voltage value. Change \(D_{IT}\) and keep iterating until the calculated \(\Delta V_G\) matches the experimentally measured shift of the gate voltage (see Fig. 17.8 (b)).

4) Therefore, the interface capacitance is obtained from: \(C_{IT} = qD_{IT}(\psi_s)\).

5) The net capacitance \((C_T)\) is then calculated using Eq. 17.4.
Method B)

The better approach is to the C-V measurement at both low and high frequency regions. The capacitance at each region is therefore given by:

\[
C_{LF}(V_G) = \frac{C_{ox}(C_B(\psi_s) + C_{IT}(\psi_s))}{C_{ox} + C_B(\psi_s) + C_{IT}(\psi_s)}
\]

\[17.5\]

\[
C_{HF}(V_G) = \frac{C_{ox}C_B(\psi_s)}{C_{ox} + C_B(\psi_s)}
\]

\[17.6\]

At high frequencies, charges trapped in the interface defects can not follow the signal; therefore there is no sign of the interface capacitance in the measured C-V curve. Consequently, \(C_{IT}\) is calculated using:
\[
\frac{C_{IT}}{q} \equiv D_{IT}(\psi_s(V_G)) = \frac{1}{q} \left( \frac{1}{C_{LF}(V_G)} - \frac{1}{C_{ox}} \right)^{-1} - \left( \frac{1}{C_{HF}(V_G)} - \frac{1}{C_{ox}} \right)^{-1}
\]

In Error! Reference source not found. (a) low and high frequency C-V curves are shown from which \(D_{IT}\) is calculated at each energy level (see Fig. 17.9 (b)).

![Fig. 17.9](image)

Fig. 17.9 (a) LF and HF characteristics of C-V. (b) The calculated \(D_{IT}\) obtained from part (a).

Inherent limitations of the C-V method

Although the C-V method is a simple technique, it cannot be used for quantitative study of surface traps, since it does not differentiate them from the bulk ones. Moreover, it is a slow measurement method; therefore relaxation would be a concern. On the other hand, in thin oxides the measurements is distorted by leakage. Finally, quantum mechanics-based interpretation of the results is not possible in C-V method.
17.3 **Method No2: The Method of Sub-threshold slope**

**General description of the method:** The method of sub-threshold slope is based on the time dependent study of the body coefficient in the diffusion dominated subthreshold region of operation of a MOSFET. In this section, the MOSFET function under subthreshold conditions is reviewed, the body coefficient is defined, and the method of subthreshold slope is then discussed. The reason behind the study of the body coefficient is that the interfacial traps induce a capacitance perturbation in the body factor, thus, its time dependent study is important.

**Subthreshold MOSFET operation conditions:** The operation of a MOSFET transistor below threshold conditions is dictated by a diffusion process and is similar to the operation of a bipolar transistor [iii]. Considering the schematic of the channel shown in Fig. 17.10 and by using the charge control approximation, one can write the drain current as:

\[
I_D = D_n \frac{Q_1 - Q_2}{L_{ch}}
\]

Therefore, considering the analogy with a bipolar transistor (with source being the emitter, gate being the base, and drain being the collector), Eq.17.8 becomes:

\[
I_D = q \frac{D_n}{L_{ch}} \left[ W \times W_{inv} \times \left( \frac{n_i^2}{N_A} e^{q\psi_s\beta} - \frac{n_i^2}{N_A} e^{q\psi_s\beta} e^{-qV_D\beta} \right) \right] \]

\[\text{Eq. 17.9}\]

which turns out to have an exponential dependence on gate voltage when the drain voltage is small:

\[
I_D \approx q \frac{D_n}{L_{ch}} W W_{inv} \frac{n_i^2}{N_A} \frac{qV_G}{e^{m\frac{k_BT}{}}} \approx I_0 e^{m\frac{k_BT}{}}
\]

\[\text{Eq. 17.10}\]

Immediately, from the slope of the \(\ln I_D(V_G)\) plot (see Fig. 17.11), one can determine the subthreshold slope which is:
\[ S \equiv \ln 10 \frac{d(V_G)}{d(lnI_D)} \]

Fig. 17.10 (a) 3D schematic illustration of a nMOSFET. The electrons flow from source to drain. (b) A graphical representation of the charge control model for calculation of the drain current. The y-direction is from source to drain. The energy barrier is reduced due to applying a positive voltage to the gate.
The Body Coefficient: The body coefficient reflects the fraction of the gate voltage \( V_G \) that is not dropped on the surface potential \( \psi_s \) and is defined as:

\[
m = \frac{V_G}{\psi_s} = \frac{C_{ox} + C_B}{C_{ox}}
\]

As mentioned before, when it comes to the nonideal case of the inclusion of surface defects, an additional capacitance \( C_{IT} \) is considered. As a result the body coefficient becomes:

\[
m = \frac{C_{ox} + C_B + C_{IT}}{C_{ox}} = \frac{C_{ox} + C_B + qN_{IT}}{C_{ox}}
\]

Now, the subthreshold slope can be related to the body coefficient and subsequently to the charge density:

\[
S = \ln 10 \frac{d(V_G)}{d(ln I_D)} = \frac{2.3m_kB T}{q} = \frac{2.3k_B T N_{IT}}{C_{ox}}
\]

Thus, we related the time evolution of defect density to the \( I_D - V_G \) curve.

Measurement of the Body Coefficient using the Subthreshold Slope method:
Assume that a small drain voltage is applied (~ 10 mV) and the \( I_D - V_G \) curve is obtained by measurement (Fig. 17.12). In this case, Eq. 17.10 could be written as for two different gate voltage values \( V_{G1} \) and \( V_{G2} \):

\[
ln I_{D1} \approx ln I_0 + \frac{qV_{G1}}{m_k B T}
\]
\[ \ln I_{D2} \approx \ln I_0 + \frac{qV_{G2}}{mk_BT} \]

By subtracting these values, the subthreshold slope can be calculated by using Eq. 17.14, from which \( m \), and subsequently \( N_{IT} \), can be obtained.

Finally it is worth mentioning that the type of the defect states (donor- vs. acceptor-like) can also be determined from the subthreshold slope method.

17.4 **Method No3: The Idlin method**

**General description of the method:** The method is based on the application of small pulse perturbations in the gate and on the observation of the effect of change on
threshold voltage. This is repeated several times. In between, the drain current is recovering to its normal value. One of the main advantages of this method over the C-V and subthreshold slope techniques is that no interruption is required during the measurements.

The experimental setup: The experimental setup (see Fig. 17.13) consists of a pulse generator unit (PGU), which provides small electrical stresses to the gate, and a source measure unit (SMU), which measures the effect on the drain current.

![Fig. 17.13 The experimental setup for Idlin](image)

The extraction of the defect density: The application of a pulsed gate voltage stress causes the NB TI defect formation. However, how is this formation measured in linear regime operation? The defect formation induces change in the threshold voltage and in the mobility. This biparametric scheme between mobility and threshold voltage can be compensating, but in a first order approximation, one can investigate the case when mobility degradation is small. In this case, the analysis proceeds as follows.

The drain current is therefore

\[
I_{D,\text{lin}} = \frac{C_{ox}W}{L} \mu (V_G - V_T) V_D
\]

from which, one can write the following relation between differential perturbations:

\[
\frac{\delta I_{D,\text{lin}}}{I_{D,\text{lin}}(t = 0)} = \frac{\delta \mu(t)}{\mu(t = 0)} - \frac{\delta V_T}{V_G - V_T(t = 0)}
\]
Assuming that the mobility perturbation is low, the change in threshold can be written as:

\[
\delta V_T(t) = -(V_G - V_T(t = 0)) \frac{\delta I_{D,lin}}{I_{D,lin}(t = 0)} \tag{17.18}
\]

In case of not small mobility degradation, we start from the definition of the transconductance:

\[
g_m = \frac{dI_{D,lin}}{dV_G} \tag{17.19}
\]

Therefore the transconductance is written as:

\[
g_m = \frac{d}{dV_G} \left[ \frac{C_{ox}W}{L} \mu(V_G - V_T)V_D \right] = \frac{d}{dV_G} \left[ \frac{C_{ox}W}{L} \mu_0 \frac{1}{1 + \theta(V_G - V_T)}(V_G - V_T)V_D \right] \tag{17.20}
\]

As a result:

\[
g_m \propto \frac{\mu^2}{\mu_0} \tag{17.21}
\]

which reflects a square root dependence of mobility on transconductance, that is:

\[
\mu \propto \sqrt{g_m} \tag{17.22}
\]

This, in turn, reflects the following differential ratio between mobility and transconductance:

\[
\frac{\delta \mu}{\mu_0} = \left( \frac{g_m}{g_{m0}} - 1 \right) \tag{17.23}
\]

The transconductance \((g_m)\) is measured by perturbing the gate voltage with time (see Error! Reference source not found. (a)). As a result, as shown in Error! Reference source not found. (b), the same behavior is observed in the drain current \((\delta I_{D,lin})\). Therefore, by using Eq. 17.23, \(\delta \mu\) can be calculated which in turn gives \(\delta V_T\) by using Eq. 17.17.
Now that we have considered both the cases of negligible/non-negligible mobility degradation, we are ready to connect the defect density with the overall picture through the charge-capacitance relation for the oxide:

\[
N_{IT}(t) = \frac{C_{ox} \delta V_T(t)}{q}
\]

17.5 Conclusions

In this lecture, we discussed three different methods for the extraction of information about the defect density in a MOSFET structure. The C-V method has strong limitations, such as its inability to differentiate bulk from interfacial traps. The subthreshold method provides with the capacity of extracting the trap density from the

Fig. 17.14 How the Idlin method provides the shift in threshold voltage. (a) Perturbing the gate voltage. (b) The resulting drain current time characteristics.
time dependence of the body factor. The Idlin method achieves the same through the time change in the threshold voltage of the device under study.

**QUESTIONS AND ANSWERS:**

1. What is wrong with only C-V measurement methods for NBTI and HCI?
   The inability of the method to distinguish between interfacial and bulk traps.

2. Why do people like to use C-V techniques? What method would you use for HCI measurement?
   They use C-V techniques because they don’t need complicated setup and they are simple to operate. One should use on-the-fly method because it provides with the bias points that drive the system in the linear regime, where HCI appears.

3. HCI do not relax. Why would you still want to use on-the-fly type methods?
   See Q2.

4. How is it possible that $\delta \mu / \mu$ is not universally negligible?
   The mobility can also be a function of threshold voltage

5. What are DSO, PGU, SMU?
   DSO stands for Digital Signal Oscilloscope, PGU stands for Pulse Generator Unit, and SMU stands for Source Measure Unit. Their function in the on-the-fly Idlin experiment is explained in the second section.

6. Why do the defect capacitance and substrate capacitance occur in parallel? Would the same happen in flash memory?
   Because the charge is partitioned in a parallel manner. In flash memories, the capacitances are in series, because the charge in the floating body sinks the field from the gate.
7. Of the methods discussed, which one is Poisson based and which one is based on continuity equation?
All of them are Poisson based. Continuity equation gets involved only in latter two (on the fly Idlin and subthreshold slope).

8. In the Idlin method, does the mobility-transconductance relation $\mu \propto \sqrt{g_m}$ depend on the threshold voltage at that time?
Yes, but in a first order approximation, the formalism described remains valid.

9. Does the subthreshold slope have a single value at a given measurement time? Is it changing between measurement time points?
Yes.

10. How long does it take to make C-V and I-V measurements?
Few minutes per sweep.

REFERENCES:

[4] “Material dependence of negative bias temperature instability (NBTI) stress and recovery in SiON p-MOSFETs”, S. Mahapatra et al. ECN Transactions 19 (2)