How to get accurate trap density measurements using charge pumping

Yuegang Zhao, Keithley Instruments, Inc.

The two most common methods used to characterize interface trap state densities in MOSFET devices are charge pumping (CP) and simultaneous C-V (the combination of high frequency and quasi-static C-V) measurement, which is typically done on MOS capacitors. As the size of the transistor scales down, thinner gate oxide is used to maintain proper gate control of the channel. This results in higher gate leakage current due to quantum tunneling of carriers through the thin gate oxide. The higher gate leakage makes characterization of interface traps more and more difficult. Quasi-static C-V becomes impractical for oxide thicknesses less than 3–4 nm. Even high frequency C-V measurement becomes a great challenge for oxides thinner than 2 nm. The CP technique has much more tolerance than the quasi-static C-V technique, so it can be used for gate oxides thinner than 2 nm, using special techniques to correct for excessive gate leakage [1].

After 40+ years of development, interface trap density in silicon dioxide gates is much less a concern than it was years ago, but SiO₂ as a gate dielectric material is approaching its physical and electrical limits [2]. The principal limitation is high leakage current due to quantum mechanical tunneling of carriers through the thin gate oxide [3]. Recently great attention has been paid to the use of high dielectric constant (high κ) materials, such as hafnium oxide (HfO₂), zirconium oxide (ZrO₂), alumina (Al₂O₃) and their silicates [4], as replacements for SiO₂ as gate dielectrics. Due to the high dielectric constants of these materials, high κ gates can be made much thicker than SiO₂ while maintaining the same gate capacitance. The result is lower leakage current—sometimes several orders of magnitude lower. Usually a very thin silicate layer forms between high κ film and Si substrate. The effect of this silicate layer on interface properties, as well as the charge trapping phenomena inside the high κ gate, is still to be understood. The CP technique becomes especially useful for characterizing interface and charge trapping phenomena, because the simultaneous C-V technique is very difficult (primarily due to the incapability of the quasi-static C-V technique at the high leakage level) for interface trap characterization.

This article will explain the basic CP technique and its variations and will describe some applications in characterizing charge trapping in high κ films.

Test procedures and variations

The basic charge-pumping technique involves measuring the substrate current while applying voltage pulses of fixed amplitude, rise time, fall time, and frequency to the gate of the transistor, with the source, drain, and body tied to ground. Figure 1 shows the connections. The two most common ways to do charge-pumping measurements are a fixed amplitude, voltage base sweep (Figure 2a) or a fixed base, variable amplitude sweep (Figure 2b). For high κ gate stack structures, the CP technique can quantify the trapped charge (Nₜₐ) as:

\[ N_{\text{t}} = \frac{I_{\text{cp}}}{qA} \]

(where \( I_{\text{cp}} \) is the measured charge-pumping current, \( q \) is the fundamental electronic charge, \( f \) is the frequency, and \( A \) is the area) since trapped charge beyond the silicon substrate/interfacial layer can be sensed [5].

In a voltage base sweep, the amplitude and period (width) of the pulse are fixed while sweeping the pulse base voltage. At each base voltage, body current can be measured and plotted against base voltage. The interface trap density (Dᵢ) can be extracted as a function of band-bending, based on this equation:

\[ D_{\text{i}} = \frac{I_{\text{cp}}}{qAf} \]

where \( \Delta \eta \) is the difference between the inversion Fermi level and the accumulation Fermi level [6].

A fixed base, variable amplitude sweep has a fixed base voltage and pulse frequency with step changes in voltage amplitude. The information obtained is similar to that extracted from a voltage base sweep. These measurements can also be performed at different frequencies to obtain a frequency response for the interface traps.

It’s relatively easy to perform CP meas-
urements and data analysis using a semiconductor characterization system (SCS) in combination with a pulse generator. Figure 1 illustrates the connections for a device under test (DUT) with one SMU and a pulse generator without a switch matrix; in Figure 3 a semiconductor switch matrix is included in the configuration.

Effects of gate leakage

The measured current \( I_m \) is the result of averaging of CP current and gate leakage current, as shown in the following. We assume here that gate capacitance is very small, so we can ignore transient currents due to gate capacitance response to pulses. This ceases to be a good assumption if the gate area of transistor under test is relatively large.

\[
I_m = I_{cp} + I_{leak(peak)} \cdot \text{dutycycle} + I_{leak(base)} \cdot (1 - \text{dutycycle})
\]

\( I_{cp} \) decreases strongly as frequency decreases, but the averaged leakage current is not frequency dependent. By subtracting \( I_m \) at very low frequency, leakage current is taken out of the measured \( I_{cp} \). Figure 4a shows \( N_{it} \) as a function of base voltage sweep at different frequencies. At lower frequency, because the leakage current is of the same order of the \( I_{cp} \), the \( N_{it} \) curve is strongly offset. Figure 4b shows the corrected \( N_{it} \) by using current measured at 300Hz (not shown in the figure). The measurement was done on an n-MOSFET with a base sweep at frequencies from 3kHz to 1MHz with the amplitude of the gate pulse fixed at 1.2V.

Frequency dependent trap density

CP measurements at different frequencies yield important information about interface trap density distribution across frequencies. This information is usually shown by plotting either the maximum of \( N_{it} \) (cumulative) or the differential of the maximum value of \( N_{it} \) as a function of frequency. Figures 5a and 5b show cumulative and differential trap distribution across frequency respectively. Maximum \( N_{it} \) was calculated from the \( N_{it} \) curve in Figure 4b after leakage correction.

Initial Charge filling

CP can be used to characterize initial stage of gate-channel interface. Figure 6 shows CP measurement on a “fresh” (never been tested) MOSFET at 1MHz. A series of consecutive CP measurements was repeated. As shown on the graph, the shape of the \( I_{cp} \) curve changes as well as the magnitude at lower biases. It then saturates after a number of repeated measurements. This indicates formation of interface traps due to electrical stress imposed by the CP measurement.

Stress-CP measurement

CP measurement is useful alone. It can also be used together with DC or AC stress to study charge trapping as well as new charge creation on the high \( \kappa \)-Si interface, as well as inside the high \( \kappa \) film. Charges injected into high \( \kappa \) film due to stress are calculated by integrating gate current through stress time. \( Q_{inj} = \int I_{leak} \cdot dt \). The advantage of stress-CP measurement over traditional stress-C-V measurement (which measures the shift of flat band voltage due to stress) and stress-IV (which measures the shift of threshold voltage due to stress) is that it can clearly distinguish existing charge trap centers with new centers created by stress [7]. Also, charge relaxation in CP measurement is usually much less than that in stress-C-V and stress-IV measurements, where trapped charges will relax between stress and the next measurement.
How to get accurate trap density measurements using charge pumping

Figure 4: (a) \( N_t \) extracted without leakage correction, amplitude of gate pulses is fixed at 1.2V. (b) \( N_t \) extracted with leakage correction.

Figure 5: (a) Example of maximum \( N_t \) as a function of frequency. Because traps responding to high frequency pulses also respond to low frequency pulses, maximum \( N_t \) at lower frequency is always larger than that at higher frequency. Maximum \( N_t \) is extracted from Figure 4b. (b) Frequency distribution of measured interface traps. It is extracted by differentiating maximum \( N_t \) with frequency from Figure 4a.

Figure 6: Initial trap filling on a “fresh” device measured from CP. The arrow shows increase of \( I_{sp} \) as CP is done repeatedly. Amplitude of gate pulses is fixed at 1.2V.
Conclusion

Charge pumping is a powerful tool for characterizing charging phenomena in high κ gate stacks. Its advantages include requiring less hardware, minimum software interface, relatively simple and easy to perform, good measurement accuracy on interface trap density, and good tolerance on gate leakage. It can characterize initial trap creation (on a fresh device) as well as new traps generated by injected charges (combining stress and CP).

References


About the author

Yuegang Zhao is a senior applications engineer with the Semiconductor Business Group of Keithley Instruments in Cleveland. He received his Masters Degree in Semiconductor Physics from the University of Wisconsin, and his B.S. in Physics from Peking University, Beijing, China. He has six years of experience in semiconductor processing and device physics.