Interface Trap response to RF Charge Pumping Measurements

G. T. Sasse and J. Schmitz

Abstract— **In this paper we will discuss the interface trap response to CP measurements at RF gate excitation. An explanation is given on how to accurately perform RF CP measurements, using an improved technique. Based on the observed response of the pumped charge per cycle with increasing frequencies a model is developed that is able to explain the observed roll-off. It is an extension to the well known classical model and it takes into account both the limited capture rates as well as a distribution of traps in the oxide.**

Index Terms— **Charge pumping, CMOS, tunneling, RF, trap response, dielectrics, characterization**

I. Introduction

THE charge pumping (CP) technique [1] is well known for its high accuracy of measuring the interface state density at the Si-SiO_2 interface of MOSFET devices and has been widely used for this purpose. With the decreasing thickness of the oxide layer in present day CMOS technologies a considerable leakage current can be seen. This leakage current can severely affect the correctness of the extracted interface state density from charge pumping data on these devices [2]. In figure 1 this effect is illustrated by comparing charge pumping data obtained on a device with 3 nm oxide thickness to data obtained on a 1.4 nm oxide device. The CP curves are obtained in a constant amplitude mode using a sinusoidal gate voltage with V_{pp}= 2V and f = 1 MHz. $V_{\rm bias}$ represents the mid-voltage level of the gate voltage signal.

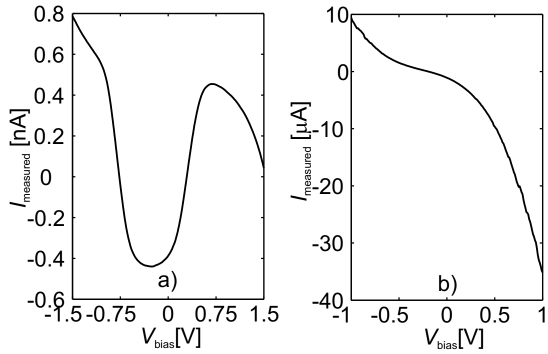

Fig. 1. Charge pumping currents obtained on n-type devices with an oxide thickness of a) 3 nm and b) 1.4 nm. The data are obtained using a sinusoidal gate voltage with V_{pp}= 2 V and f = 1 MHz. The charge pumping current is completely overwhelmed by the leakage current on the 1.4 nm oxide.

Manuscript received September 26, 2006. This work was supported by the Dutch Technology Foundation (STW). G.T Sasse and J. Schmitz are with the MESA+ Institute for Nanotechnology, Group of Semiconductor Components,University of Twente, The Netherlands (phone: +31-534894394, fax: +31-534892644, e-mail: g.t.sasse@utwente.nl).

II. RF CP Technique

As we have shown in [3] a CP current can be observed at RF large-signal gate bias, allowing interface state probing even on very leaky dielectrics. The measurement setup used for RF charge pumping is shown in figure 2. During a CP measurement a device is repeatedly switched between inversion and accumulation. During inversion charge carriers in the channel region get trapped into the interface states and they are released towards the substrate during accumulation. This leads to a net amount of charge being transferred from the substrate through the interface states towards the source/drain. Repeatedly switching the device causes this charge transfer to occur repeatedly as well and thereby a DC current can be observed at either the substrate or the drain/source contact. This current is the charge pumping current $I_{\rm cp}$. $I_{\rm cp}$ is frequency dependent, because with increasing frequency the charge pumping cycle occurs more often, thereby pumping more charge from substrate towards source/drain per unit time. The idea of the RF charge pumping technique is to to increase the CP component by making use of this frequency dependence of $I_{\rm cp}$. Using gate voltage signals with frequencies into the GHz range, the CP current can be so large that the large tunneling component as observed in figure 1 can completely be eliminated. This is shown in figures 3 and 4. In these figures RF CP results are shown of a 1.4 nm oxide device, which suffers from the high leakage current as shown in figure 1b. The test structures used in these measurements are devices optimized for two-port RF measurements with small channel length and source and drain tied together. The CP current is measured at the drain/source terminal. From figures 3 and 4 it is clear that, for this device, useful CP data can only be obtained at frequencies larger than 100 MHz, a frequency range not covered by the classical CP approach, thereby illustrating the value of the RF CP technique.

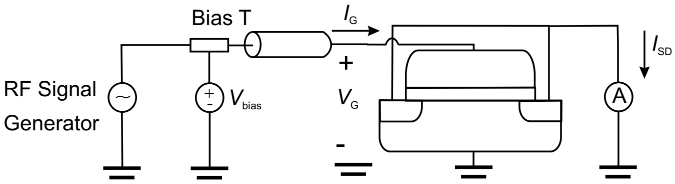

Fig. 2. Schematic drawing of the RF CP measurement setup.

III. RF gate Voltage Generation

Generating well defined RF gate voltage signals is not straightforward, due to reflections occurring at the gate. In [3] we have shown that the harmonic distortion introduced on

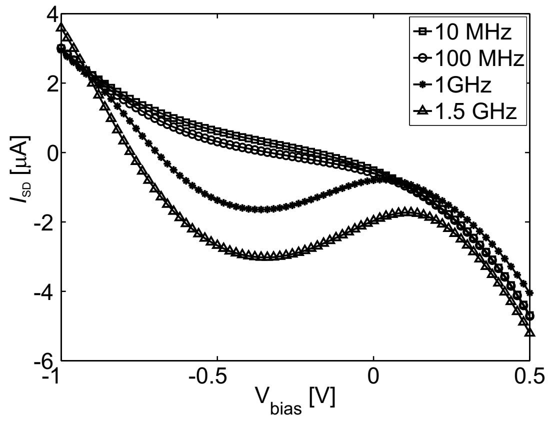

Fig. 3. Measured charge pumping characteristics on a device with 1.4 nm oxide thickness, for different frequencies. The input power is set so that V_{PP} = 1.8 V.

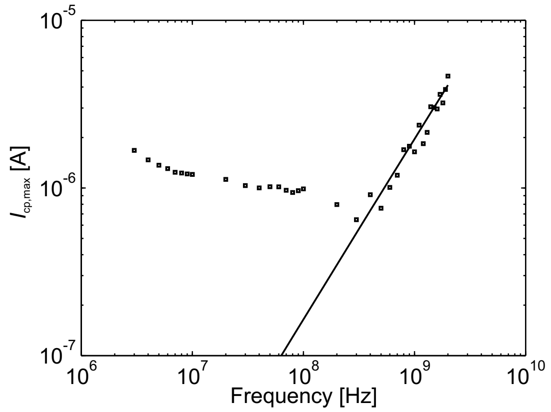

Fig. 4. Maximum CP current plotted versus frequency on a 1.4 nm oxide device. The applied gate voltage level V_{PP} is 1.8 V. The solid line represents the frequency response as expected from theory.

our test structures is negligible for frequencies up to 2 GHz when using sinusoidal voltage signals. The technique of setting the desired gate voltage level by calculating transmission line equations in the time domain [3], proved not to be accurate enough, resulting in the fluctuating data points above 100 MHz as shown in figure 8a. We developed a more accurate technique of obtaining the desired amplitude. This technique makes use of the nonlinear relation between the gate voltage and the gate leakage current. Due to the time-varying gate voltage signal a DC current I_{DC} can be observed, which is defined by:

$$I_{\mathrm{DC}} \equiv \frac{1}{T}\int_0^T I_{\mathrm{gate}}(t)dt \qquad (1)$$

This I_{DC} is basically a measure for the net amount of charge flowing through the gate per unit time. It is important to note that only the tunneling component of the gate current contributes to this I_{DC} as a capacitive current does not result in a net amount of charge transfer over a period of the gate voltage signal. Furthermore if the DC bias voltage V_{bias} is chosen properly, the CP effect does not occur. The fact that only the tunneling current contributes to I_{DC} results in the fact that I_{DC} is a frequency independent parameter. Besides this, I_{DC} is also very sensitive to the amplitude of the sinusoidal gate voltage signal. This makes it a very good indicator for the realized gate voltage amplitude. This effect is illustrated in figures 5, 6 and 7. In figure 5 we have plotted the measured

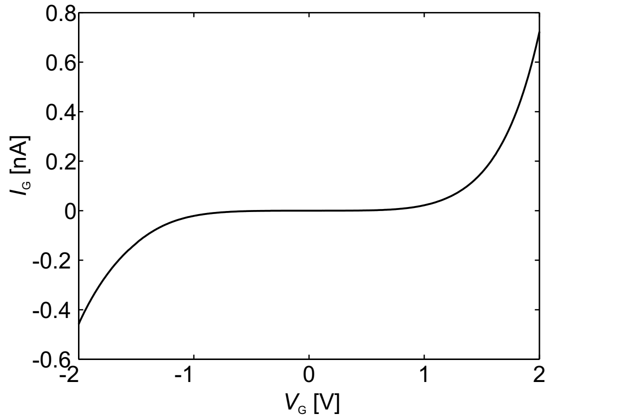

Fig. 5. Measured gate current characteristic. The current is measured at the DC node of the bias T of figure 2.

DC leakage current as a function of gate voltage. Using the current-voltage relation we can obtain the result as shown in figure 6, where both the gate voltage as well as the leakage current is plotted as a function of time. This result is calculated using a gate voltage signal which is given by:

$$V_{\mathrm{G}}(t) = V_{\mathrm{bias}} + \frac{V_{\mathrm{PP}}}{2} sin\,(2\pi f t) \qquad (2)$$

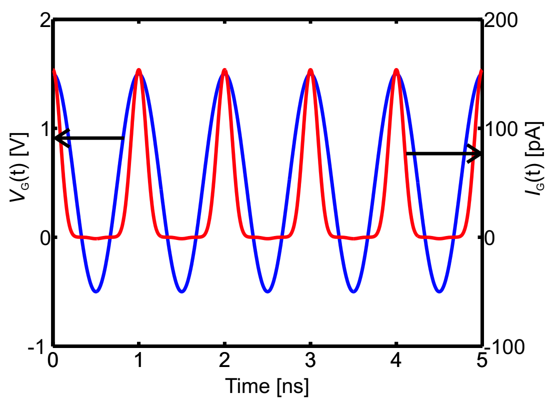

Fig. 6. Leakage current as a function of time, obtained using the results of figure 5. The corresponding gate voltage signal is also shown.

In the calculation of the result in figure 6 we have used $V_{\mathrm{bias}} = 0.5V$ and $V_{\mathrm{PP}} = 2V$. In order to illustrate how we can make use of this leakage current for determining the gate voltage amplitude, we plotted the total net amount of charge flowing through the gate as a function of time. The total amount of charge is basically the integral of the current over time. This is shown in figure 7, for three different frequencies and three different values of V_{PP}. The associated values of I_{DC} that are given besides the figure are obtained by taking the derivative of Q_{tunnel} to time. We see that this derivative is completely frequency independent and hence the values of I_{DC} are identical for all three different frequencies used. Furthermore we can recognise the strong dependence of I_{DC}

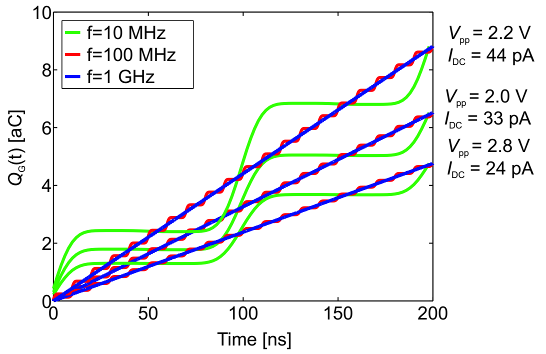

Fig. 7. Net amount of charge flowing through the gate as a function of time. The result is obtained from calculations as the one shown in figure 6, based on the measured leakage current characteristic of figure 5. The results are obtained at three different frequencies and three values of Vpp. The result clearly shows the frequency-independence of I_{DC} as well as the strong dependence on gate voltage amplitude.

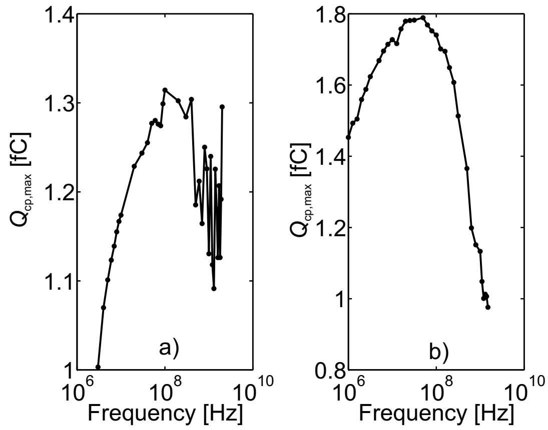

Fig. 8. Pumped charge per cycle obtained a) using the technique as explained in [3] and b) using I_{DC} . The data are obtained on two different devices.

on the amplitude of the gate voltage signal From this we can conclude that using I_{DC} we can carefully set the amplitude of the sinusoidal gate voltage signal, for frequencies into the GHz range. An important issue with this technique is the fact that the gate leakage current cannot directly be measured at the gate contact. This can be understood by looking at figure 2: the bias T prevents a proper measurement of I_{DC} at this terminal. Therefore we measure I_{DC} at the drain/source contact. In order to come to an accurate result we make sure that the device is constantly biased in inversion, or in other words $V_{\mathrm{bias}} - \frac{V_{\mathrm{pp}}}{2} > V_{\mathrm{T}}$. Now a low-ohmic channel region is constantly present, which causes the gate leakage current to flow completely to the drain/source terminal. In practice the I_{DC} technique works as follows: First we generate a gate voltage signal with the desired amplitude at a frequency of 1 MHz. At this frequency the effect of standing waves on the measurement cables can be ignored and the amplitude of the voltage signal at device level can be accurately measured using an oscilloscope. The associated value of I_{DC} is recorded. Subsequently, in order to generate an RF voltage signal with the same amplitude, we connect an RF power source and set the frequency to the desired value. Now the available power of the RF power source can be very gradually increased while I_{DC} is constantly monitored. We stop this process until the desired I_{DC} (as obtained at 1 MHz) has been reached. The power level at which this happens is the appropriate power level for setting the amplitude of the RF voltage signal to the desired level at the input of the device under test (on wafer). The impact of this technique w.r.t the technique as used in [3] is shown in 8. In this figure we have plotted the pumped charge per cycle against frequency where the voltage signals are generated using the technique as in [3] and the I_{DC} technique. Using the I_{DC} technique we are now able to analyze the interface state behavior at frequencies of more than two orders of magnitude higher than using the conventional CP technique.

IV. Physical Model of Trap Response

Looking at the result of figure 8b, we see that the pumped charge per cycle decreases with frequency above $\sim$ 100 MHz. In order to explain this roll-off, we look at the physical process occurring during charge pumping. First we look at the classical theory of charge pumping [1]. From this theory we know that that Q_{cp} is given by:

$$Q_{cp} = qA_{\mathrm{G}}\overline{D_{\mathrm{it}}}\Delta E \tag{3}$$

In this expression q is the elementary charge and A_{G} is the surface area of the device. $\overline{D_{\mathrm{it}}}$ is the effective interface state density and ΔE is the energy window between which interface states are located that contribute to the CP effect. This ΔE is given by [1] :

$$\Delta E = E_{\mathrm{em,e}} - E_{\mathrm{em,h}} \tag{4}$$

In this expression $E_{\mathrm{em,e}}$ represents the highest energy level for which the nonsteady state emission process of electrons is negligible. $E_{\mathrm{em,h}}$ is the lowest energy level for which the nonsteady state emission process of holes is negligible. Expressions for $E_{\mathrm{em,e}}$ and $E_{\mathrm{em,h}}$ can be found directly from the theory of [1] and they are given by:

$$E_{\mathrm{em,e}} - E_{\mathrm{i}} = -kTln\left(v_{\mathrm{th}}n_i t_{\mathrm{em,e}}\sigma_{\mathrm{n,p}} + e^{\frac{E_{\mathrm{i}}-E_{\mathrm{F,inv}}}{kT}}\right) \tag{5}$$

$$E_{\mathrm{em,h}} - E_{\mathrm{i}} = kTln\left(v_{\mathrm{th}}n_{\mathrm{i}} t_{\mathrm{em,h}}\sigma_{\mathrm{n,p}} + e^{\frac{E_{\mathrm{F,acc}}-E_{\mathrm{i}}}{kT}}\right) \tag{6}$$

In these expressions we recognise $t_{\mathrm{em,e}}$ and $t_{\mathrm{em,h}}$, these are the times available for the nonsteady-state emission of electrons and holes respectively. For sinusoidal gate voltage signal these can be found using the expressions given in [4]. At increasing frequencies these times decrease, thereby broadening ΔE. This results in an increase in $Q_{\mathrm{cp,max}}$ with frequency as can be observed with CP data at low frequencies. Furthermore in expressions 5 and 6 we see energy levels $E_{\mathrm{F,inv}}$ and$E_{\mathrm{F,acc}}$; these are the energy levels occurring at inversion and accumulation conditions respectively. The maximum possible ΔE is equal to $E_{\mathrm{F,inv}}$ - $E_{\mathrm{F,acc}}$, thereby stopping the effect of increasing $Q_{\mathrm{cp,max}}$ with increasing frequency for very high frequencies. In figure 9 we show a best fit of expression 3 with the measured data points in the line depicted "classical". We do recognise a decrease in the slope of $Q_{\mathrm{cp,max}}$ w.r.t. frequency, but the roll-off that can be seen above 100 MHz cannot be explained using the classical CP theory.

In order find a possible explanation of the observed roll-off, we take into account the incomplete filling of the states during inversion and accumulation conditions. This incomplete filling of the states is a result of the limited capture rates of charge carriers into the interface states. At increasing frequencies the times available for these capture processes are so small that not all carriers will respond to the gate voltage signal; this means that a reduced number of interface states is active during charge pumping, leading to a reduction in $Q_{cp,max}$. If we assume that the interface state population can be described using one value for both the interface state density and the capture cross section of the traps, we can derive an expression for the difference in trap occupancy level Δf_{T} between inversion and accumulation conditions. In [5] an expression was found for Δf_{T} when using trapezoidal gate voltage signal. It can be reduced to an expression with equal parameters for the capture process of electrons and holes under maximum CP conditions [6]. When using sinusoidal gate voltage signals this filling function needs to be adjusted to take into account the time-dependent capture rates. We do this by making use of the time that the device is actually biased in inversion respectively accumulation. This time is inversely proportional to the gate voltage frequency. Now we can derive an expression for Δf_{T} given by:

$$\Delta f_{\mathrm{T}} = \frac{\left(1 - e^{\frac{-c_{\mathrm{n,p,max}}}{Z_{\mathrm{c}} f}}\right)^2}{\left(1 - e^{\frac{-2c_{\mathrm{n,p,max}}}{Z_{\mathrm{c}} f}}\right)} \tag{7}$$

In this expression $c_{\mathrm{n,p,max}}$ is the maximum capture rate over one cycle and parameter Z_{c} relates the gate voltage to the effective capture rate. We can include the effect of the incomplete filling of the states by multiplying the classical expression for the CP current with expression 7. The resulting frequency response of $Q_{\mathrm{cp,max}}$ is shown in 9, by the line depicted "single capture rate". This line is a best fit to the measured data , where $\overline{D_{\mathrm{it}}}$ $c_{\mathrm{n,p,max}}$ and Z_{c} are used as fitting parameters. It is clear that expression 7 is still not able to exactly describe the observed frequency response of $Q_{\mathrm{cp,max}}$. The roll-off is much steeper than the roll-off observe from the measured data.

This observation leads to the assumption that it is not correct to model the interface state population as one homogeneous group with only a single parameter $\overline{D_{\mathrm{it}}}$ and $\sigma_{\mathrm{n,p}}$. As already observed from CP measurements in small voltage mode [6], the trap distribution can be described using an exponential function as:

$$N_{\mathrm{T}}(K) = N_0 e^{-\frac{K}{\kappa\alpha}} \tag{8}$$

In this expression K is the distance a trap is located away from the interface, κ is the characteristic parameter of the capture process as in [5] and parameter α is a fitting parameter. The pumped charge per cycle $Q_{\mathrm{cp,max}}$ can be found by integrating the contribution of all traps over distance and approximating Δf_{T} with a step function around K_{max}, as was done in [6]. K_{max} is defined as the value of K where $\Delta f_{\mathrm{T}} = 0.5$ and can be found from expression 7, making using of the dependence of $c_{\mathrm{n,p,max}}$ on the capture cross section $\sigma_{\mathrm{n,p}}$ and hence on K. This results in:

$$Q_{\mathrm{cp,max}} = \int_0^{K_{\mathrm{max}}} -2qA_{\mathrm{G}}N_{\mathrm{T}}(K)ln\left(v_{\mathrm{th}}n_{\mathrm{i}}t_{\mathrm{em}}\sigma_0 e^{-\frac{K}{\kappa}}\right) dK \tag{9}$$

The integrand in this expression is similar to the "classical" expression. The exponential term describes the dependence of $\sigma_{\mathrm{n,p}}$ on K; t_{em} is the time available for the non-steady-state emission of electrons and holes. The "distributed traps" curve of figure 9 is calculated using expression 9 with $\kappa \cdot N_0 = 2.23 \cdot 10^{10} eV^{-1} cm^{-2}$ and $\alpha = 2.2$. The "distributed traps" model completely describes $Q_{\mathrm{cp,max}}$ accurately over the entire frequency range. The capture cross section of traps located at $K = 0$,σ_0 is 6.58·10^{-15} cm^2, this is a realistic value. Furthermore if we calculate the total numbers of traps in the oxide from the distribution equation 8 we come to a total interface state density of 4.9·10^{10} eV^{-1}cm^{-2}. In the fit of the classical curve and the "single capture rate" we used an interface state density of 4.8·10^{10} eV^{-1}cm^{-2}. This is in very good accordance with the new model.

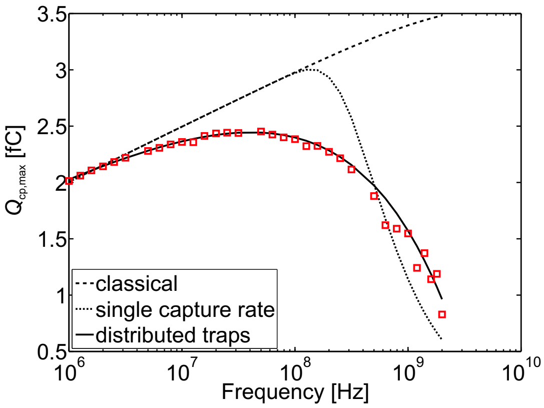

Fig. 9. Measured pumped charge per cycle on a 3nm oxide device with an applied gate voltage level V_{pp} of 2 V. The red squares represent the data points and the three lines are theoretical curves using the three different models.

V. Discussion

As already observed in earlier works (see e.g. [6]), an increase in the measurement frequency may lead to a roll-off in measured Qcp,max with frequency. If one wants to extract the interface state density on a very leaky dielectric as that of figures 3 and 4, the only available data is found at frequencies where roll-off has already set in. Assuming an exponential distribution of the traps and with parameters $c_{\mathrm{n,p,max}}$ and Z_{c} known, it is possible to find the interface state density by fitting the distribution curve and parameter σ_0 to the measured data. In this work we have used $c_{\mathrm{n,p,max}}$ and Z_{c} as fitting parameters as well. Furthermore, from the RF CP data the number of traps that are actually active at these frequencies can be found. This number of active traps has a strong dependence on the surface potential. In order to analyze the RF behavior of interface states for a given device, one could perform RF CP measurements at different gate voltage levels, ranging from the small voltage levels used in e.g. [6] to the conventional large signal swing.

VI. Conclusions

In this paper we have shown that using the RF CP technique, charge pumping data can be obtained at frequencies of up to 2 GHz. An improved procedure for obtaining the desired gate voltage levels is presented. The observed roll-off of the pumped charge per cycle at frequencies above 100 MHz can completely be explained by the incomplete filling of traps and assuming an exponential distribution of the traps in the oxide. The RF CP technique allows obtaining the interface state density on leaky dielectrics, and can be employed to fully characterize the trap behavior at radio frequencies.

Acknowledgments

We would like to thank Philips Research for the use of their samples.

References

[1] G. Groeseneken, H.E. Maes, N. Beltrán and R.F. De Keersmacker, "A Reliable Approach to Charge-Pumping Measurements in MOS Transistors", *IEEE Trans. Electron Devices*, vol. ED-31, p. 42, 1984.

[2] P. Masson, J-L. Autran and J. Brini, "On the Tunneling Component of Charge Pumping Current in Ultrathin Gate Oxide MOSFET's", *IEEE Electron Device Lett.*, vol. 20, p. 92, 1999.

[3] G.T. Sasse, H. de Vries and J. Schmitz, "Charge Pumping at Radio Frequencies", in *Proc. ICMTS*, 2005.

[4] J.L. Autran and C. Chabrerie, "Use of the charge pumping Technique with a sinusoidal gate waveform", *Solid-State Electron.*, vol. 39, p. 1394, 1996.

[5] D. Bauza and G. Ghibaudo, "Analytical study of the contribution of fast and slow oxide traps to the charge pumping current in MOS structures", *Solid-State Electron.*, vol. 39, p. 563, 1996.

[6] Y. Manéglia, F. Rahmoune and D. Bauza, "On the Si-SiO_2 interface trap time constant distribution in metal-oxide semiconductor transistors", *J. Appl. Phys.*, vol.97, p. 014502, 2004.