A Novel and Direct Determination of the Interface Traps in Sub-100nm CMOS Devices with Direct Tunneling Regime (12~16A) Gate Oxide

Steve S. Chung[1], S.-J. Chen[1], C.-K. Yang[2], S.-M. Cheng[2], S.-H. Lin[2], Y.-C. Sheng[2], H.-S. Lin[2], K.-T. Hung[2], D.-Y. Wu[2], T.-R. Yew[2], S.-C. Chien[2], F.-T. Liou[2], and Frank Wen[2]

[1]*Department of Electronic Engineering, National Chiao Tung University, Taiwan*

[2]*UMC, Science-based Industrial Park, Hsinchu, Taiwan*

Abstract- For the first time, an improved charge pumping (CP) method has been implemented for direct determination of the interface traps in ultra-short gate length CMOS devices with ultra-thin gate oxide in the direct tunneling regime. The leakage current in a 12-16A gate oxide can be removed from the measured CP current, which enables accurate determination of the interface traps. This method has been demonstrated successfully for various RTNO grown and RPN treated oxide CMOS devices with very thin gate oxide. Moreover, it can be used as a good monitor of ultra-thin gate oxide process and the evaluations of device reliabilities in relating to the interface trap generation.

Introduction- As predicted from the SIA roadmap, scaling of sub-100nm device in 2005 will need a t_{ox} in the range 10-15A. One faces a big problem of how to measure the oxide quality with thickness 10 to 20A, in particular the interface traps, N_{it}. So far, there is no definite approach to measure the N_{it} for very small and ultra-thin gate oxide devices.

It is also well known that two pronounced effects occur as a result of the gate oxide scaling below 30A, i.e., direct tunneling gate leakage and the quantum mechanical effect [1-2]. This makes the device characterization more difficult. Basically, two methods can be employed to determine the interface traps in a CMOS device, i. e., conventional CV method [3] and the CP method [4-7]. However, CV method needs a large area capacitor for the measurement and will encounter difficulties when gate leakage current exists. Until recently, the CP method on a thin gate oxide measurement has aroused much interest [8]. However, this method is in-accurate when $t_{ox} \leq 12$A.

In this paper, two new methods derived from the conventional charge pumping (CP) method have been demonstrated for accurate determination of the interface traps in CMOS devices with short channel length (very small dimension) and ultra-thin gate oxide.

I. Device Preparation

The devices used in this study were fabricated by the state-of-the-art IC manufacturing. 12-16A gate oxides were formed by rapid thermal nitric oxide (RTNO). And, remote plasma nitridation (RPN) treatment was used for a split of 16A devices after gate oxide formation to reduce 2-3 current orders of gate leakage. The masked lengths ranged from 0.22 to 0.11um (gate ashing length ≈ 0.04um, to form nano-meter channel length device) were used.

II. Results and Discussion

A. Principle of the Method and Experimental Observations

Figure 1 shows the popular schematic of the two-side [4-6] and single-side [7] CP measurements. In method 1, with both S/D grounded and by applying a gate pulse with fixed base voltage (V_{gl}), the channel will operate between accumulation and inversion. This gives rise to the charge pumping current I_{CP} measured from the bulk. Method 2 has floating source and the I_{CP} is measured from the drain. As we see from Fig. 1(b), the leakage of I_{CP} is very small at low V_{gh} when $t_{ox} > 30$A. However, it is noticed that the leakage current becomes dominant for t_{ox} less than 20A.

B. Comparison of the Leakage for Two CP Setups

Figs. 2 and 3 show that Method 2 is not adequate for the CP measurement since the leakage current becomes dominant in the accumulation region. While, method 1 is well suited for the CP measurement purpose with suitable chosen gate voltage.

C. New Charge Pumping Methodology

Fig. 1(c) shows the steps of the new method. Experimental results in Figs. 4 and 5 show that: (1) the bulk current decreases with reducing channel length. This means that it is better to measure a low leakage CP current with a shorter channel length device, and (2) the leakage current increases with reducing t_{ox} (at $V_G < 0$V). Before calculating N_{it} from I_{CP}, we need to remove the leakage current from the I_{CP}. Here, we provide two different methods to achieve this purpose.

(1) *High-low frequency CP method* (modified from [8])- First, we measure the I_{CP} for various frequencies as in Fig. 6. At low frequency, group 2 curves are considered as the leakage current. Curve 1 is the I_{CP} at high frequency. Curve 1 subtracting group 2 curves gives the correct I_{CP}(group 3) without leakage component.
(2) *Incremental frequency CP method-* From the measured I_{CP} for various frequencies, we take the difference of I_{CP} between two successive frequencies as shown in Fig. 7. For example, $I_{CP(1MHz)}$-$I_{CP(500kHz)}$ is regarded as the I_{CP} at 500kHz since I_{CP} is proportional to f.

Both methods give a close result of I_{cp} for 1MHz signal (e.g., compare curve 3 in Fig. 6 and curve A in Fig. 7). Since the leakage component is close at two successive frequencies, the *incremental frequency CP method* can give more accurate results. As expected, even for very-thin ($t_{ox} \leq 12$A) gate oxide devices, this new CP methodology is still valid.

D. Determination of the Interface Traps and Applications

To determine the interface traps, N_{it} can be calculated from the $I_{CP,MAX}$. The equations, method, and results are given in Table I, Figs. 8 and 9, respectively. Fig. 9 shows the calculated N_{it} per unit width for the measured 80 devices with n- and p-channel. It is noted that : (1) a thicker gate oxide exhibits larger N_{it} as a result of a longer thermal treatment, (2) RPN treated gate oxide has larger N_{it}, and (3) the slopes of these curves give the N_{it} values, which can be used as a monitor of the oxide quality. In short, the developed method not only can be used to calculate the N_{it} values but also be useful as a monitor of the oxide quality in an ultra-thin gate oxide process.

In summary, a new CP methodology, based on two different CP methods, has been developed for ultra-short channel length and ultra-thin gate oxide in the range 12-16A. It allows fast and easy calculation of the N_{it} generated during the process. This method is superior to the conventional CV method for N_{it} characterization in that the latter needs a large area capacitor samples. On the other hand, both methods can also be applied to the device hot carrier reliability in relating to the generated interface traps for nano-scale devices.

 0-7803-7312-X/02/$17.00 ©2002 IEEE

Acknowledgments- This work was supported in part by the National Science Council, Taiwan, under contract NSC89-2215-E009-107.

References

[1] C.-H. Choi et al., *Symposium on VLSI Tech.*, pp. 63-64, 1999.

[2] K. Yang et al., *Symposium on VLSI Tech.*, pp. 77-78, 1999.

[3] Lewis M. Terman, *Solid-State Electronics*, Vol. 5(5), p. 285-299, 1962.

[4] S. S. Chung et al., *Proc. SSDM*, pp. 841-843, 1993.

[5] S. S. Chung et al., *IEEE T-ED*, vol. 46, pp. 1371-1377, 1999.

[6] P. Heremans et al., *IEEE T-ED*, Vol. 36, pp. 1318-1335, 1989.

[7] C. Chen et al., *IEEE T-ED*, Vol. 45, No. 2, pp. 512-519, 1998.

[8] P. Masson et al., *IEEE EDL*, Vol. 20, No. 2, pp. 92-94, 1999.

(1a) $L_{MASK} = L_{gate} + 2 * \frac{\Delta L_1}{2} = L_{gate} + \Delta L_1$

(1b) $L_{gate} = L_{eff} + 2 * \frac{\Delta L_2}{2} = L_{eff} + \Delta L_2$

(1c) $\Delta L_0 \cong \Delta L_1 + \Delta L_2$

(2a) $N_{it,1,total} = N_{it,11} + N_{it,12}$

(2b) $N_{it,2,total} = N_{it,21} + N_{it,22}$

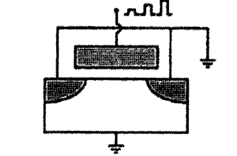
(2c) $\Delta I_{CP,max} \propto \Delta N_{it,total} = N_{it,1,total} - N_{it,2,total}$

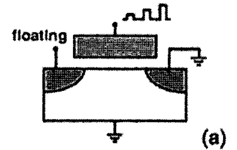
$= (N_{it,11} + N_{it,12}) - (N_{it,21} + N_{it,22})$

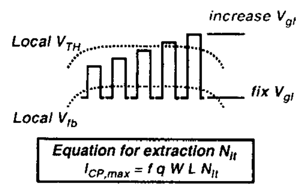
(if $N_{it,11} \approx N_{it,21}$)

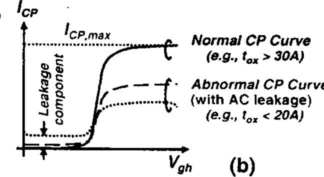
$= N_{it,12} - N_{it,22} \propto \Delta L$

Table I Extraction of offset length ($\Delta L_0 \approx \Delta L_1$ (ashing length) + ΔL_2 (overlap region)) from interface traps distribution.









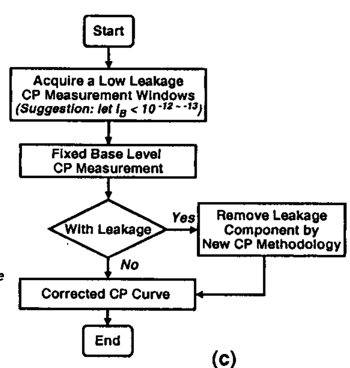


Fig. 1 (a) Two-side and single-side charge pumping (CP) setups. (b) *Upper-* Local threshold voltage (V_{TH}) and flat band (V_{fb}) distribution in relating to low level and high level gate voltage (V_{gl} and V_{gh}). *Bottom-* Normal and abnormal CP curves. (c) The flow chart of the new CP methodology.

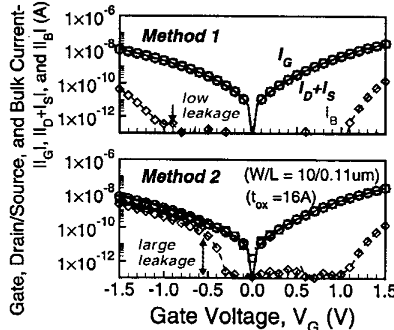


Fig. 2 Current components for two-side and single- side CP setups. Note that: (1) $I_G \approx I_D + I_S$; (2) Method 1 has lower bulk current and hence has lower leakage component for CP measurement.

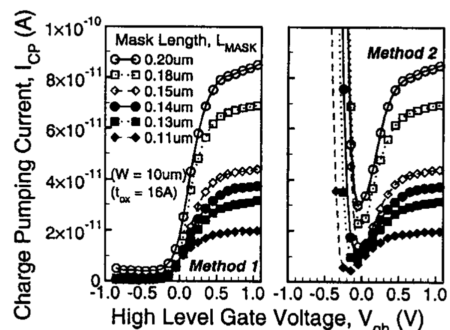


Fig. 3 Two-side and single-side CP curves. Method 2 shows a huge leakage current at low V_{gh} (accumulation region).

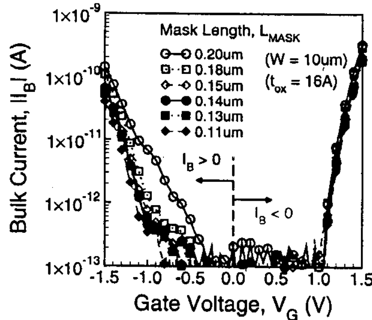


Fig. 4 Length dependent bulk currents. Shorter channel device has smaller bulk current.

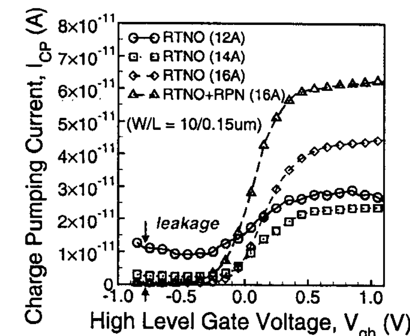


Fig. 5 Measured CP currents for ultra thin (12~16A) gate oxide. Note that 12A gate oxide has large leakage currents for V_{gh} < 0V.

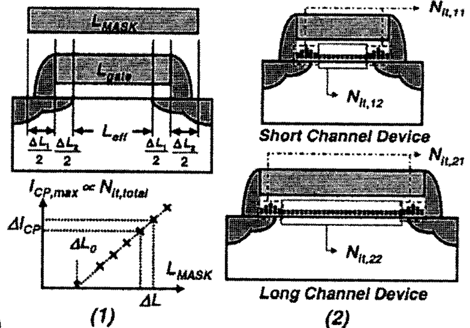


Fig. 8 Illustration of ΔL_0 extraction from CP data. (1) Parameter definition and extraction method. (2) Interface traps distribution in short and long channel length devices.

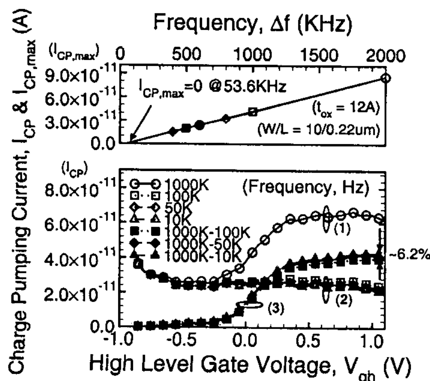


Fig. 6 *Bottom-* High-low frequency CP method (i.e., $I_{CP(1MHz)} - I_{CP(10KHz)} \approx I_{CP(1MHz)}$). *Upper-* Frequency dependent maximum CP currents.

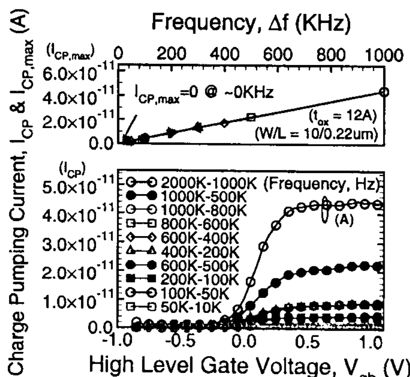


Fig. 7 *Bottom-* Incremental frequency CP method (i.e., $I_{CP(2MHz)} - I_{CP(1MHz)} \approx I_{CP(1MHz)}$). *Upper-* Frequency dependent maximum CP currents.

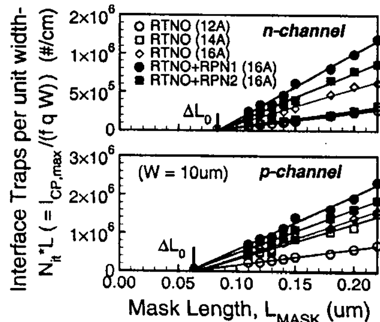


Fig. 9 Calculated N_{it} from $I_{CP,MAX}$ in Fig. 7. It also shows the extraction of offset length ΔL_0 (≈ ΔL_1 (≈ 0.04um) + ΔL_2).