A novel modified charge pumping method for trapped charge characterization in nanometer-scale devices*

Zhu Peng(祝鹏)[1,†], Pan Liyang(潘立阳)[1,2], Gu Haiming(古海明)[1], Qiao Fengying(谯凤英)[1], Deng Ning(邓宁)[1,2], and Xu Jun(许军)[1,2]

(1 Institute of Microelectronics, Tsinghua University, Beijing 100084, China)
(2 Tsinghua National Laboratory for Information Science and Technology, Beijing 100084, China)

Abstract: A new modified method based on the charge pumping technique is proposed and adopted to extract the lateral profiles of oxide charges in an advanced MOSFET. A 0.12 μm SONOS device with 50 nm threshold voltage peak is designed and utilized to demonstrate the proposed method. The trapped charge distribution with a narrow peak can be precisely characterized with this method, which shows good consistency with the measured threshold voltage.

Key words: charge pumping; trapped charge distribution; localized V_{T}
DOI: 10.1088/1674-4926/31/10/104008 **PACC:** 7220J; 7340Q

1. Introduction

As transistor gate lengths are scaled towards the deep submicron scale, hot carrier degradation of transistor characteristics forms a more severe limitation on device operation in MOS transistors[1−3]. This local degradation results from two factors: interface state generation and charge trapping in the gate insulator[4]. To understand and model this degradation, it is essential to study the spatial distribution of the trapped charges, which are scaled into the nanometer-scale as well.

The charge pumping (CP) technique is well known for its high sensitivity and accuracy of measuring the interface state density at the Si–SiO_2 interface[5] and extracting the localized trapped charge profile in the gate oxide[6−8]. Although satisfactory results have been obtained using this technique for long channel devices, it was found that as the effective channel length decreases, the accepted charge pumping method provides an imprecise evaluation of the trapped charge distribution. The loss in accuracy is due to the narrow localized threshold voltage (V_{T}) peak across the channel caused by source/drain proximity or trapped charges in the gate dielectric.

In this paper, we analyze the effect of a narrow V_{T} peak on the precision of the traditional CP technique and propose a new modified CP method to obtain a more precise distribution of trapped charges in a scaled MOSFET. Moreover, a silicon–oxide–nitride–oxide–silicon (SONOS) device, which has localized charge trapping in the nitride layer and better controllability of the width and density of the trapped charges by accurate transient program operation, is appropriate to be utilized to investigate the charge trapping characteristics. A 0.12 μm SONOS device, with the Fowler–Nordheim (FN) tunneling program and band-to-band hot hole (BBHH) injection erase which may cause the V_{T} peak to be narrower than 50 nm, is designed to evaluate the accuracy of the traditional technique and the proposed CP method.

2. Method

A reliable technique based on charge pumping was reported by Chen and Ma in 1996 which has been widely used to extract the lateral profile of oxide charge[8]. The CP measurement is realized by applying a repetitive square pulse with a fixed base voltage level (V_{b}) and a varying high voltage level (V_{h}) to the gate of the transistor, which is the so-called "constant V_{b}" ($\mathrm{CV_b}$) method. The voltage signal applied to the gate will alternately accumulate and invert the channel. Therefore, traps at the Si–SiO_2 interface will be charged and discharged. The involved electrons and holes will flow back and forth at source/drain and substrate respectively, forming the so-called charge pumping current (I_{cp}) at these terminals[9].

The localized V_{T} along the channel is related to doping concentration and trapped charges. I_{cp}–V_{h} curves of the MOSFET in virgin and charge trapped states are measured from the drain or source junction (with the other junction floating). The I_{cp} curve on an undamaged MOSFET provides the threshold voltage ($V_{\mathrm{T}(x)}$) and flat-band voltage ($V_{\mathrm{FB}(x)}$) distribution. Since the main effect of the trapped charges is to shift V_{T} and V_{FB} (in the form of V_{h}), the localized V_{T} is then used to obtain the trapped charge distribution. It is worth noting that the CP method can also be utilized to profile the distribution of interface states. However, interface state generation after stress combined with trapped charges make the characterization more complicated. In such a case, a mild neutralization step in CP testing is commonly needed to neutralize the trapped charges and it should be ensured that this step does not change the interface state density[8]. This paper focuses on the effect of trapped charges and assumes that the interface states are uniformly distributed for simplicity. Therefore, the equations used in $\mathrm{CV_b}$ to calculate the N_{ox} versus x curve are as follows:

$$x = L\frac{I_{\mathrm{cp}}(V_{\mathrm{h}})}{I_{\mathrm{cp,\,max}}}, \tag{1}$$

* Project supported by the National Basic Research Program of China (No. 2006CB302700), the National Natural Science Foundation of China (No. 60876076), and the National Key Scientific and Technological Project of China (No. 2009ZX02023-5-3).
† Corresponding author. Email: zhup07@hotmail.com
Received 2 April 2010, revised manuscript received 4 May 2010

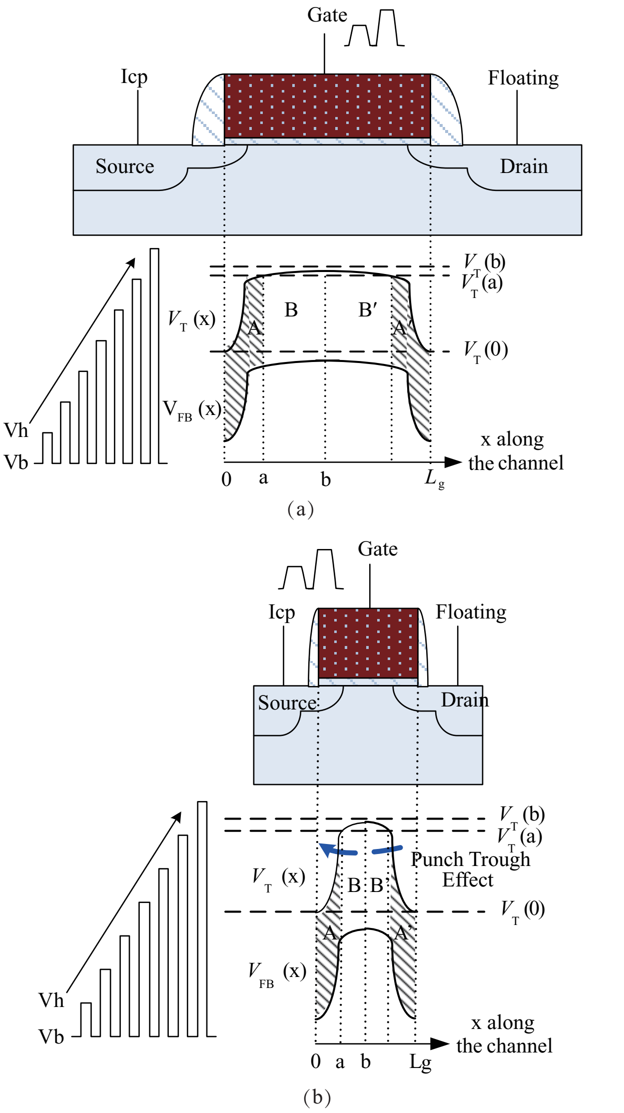

Fig. 1. Illustration of the CV_b method in (a) a long channel and (b) a short channel MOSFET.

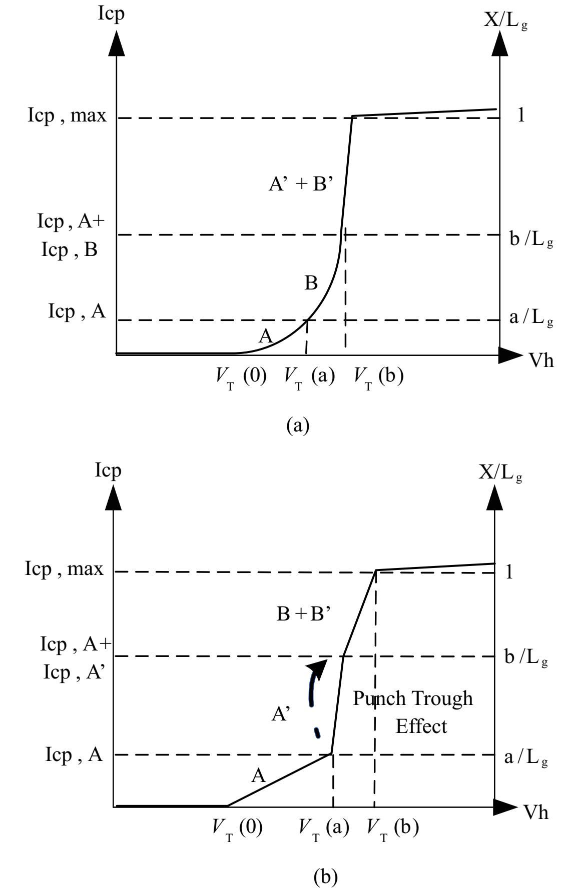

Fig. 2. Illustration of the I_{cp} curve method in (a) a long channel and (b) a short channel MOSFET.

$$N_{ox}(x) = \frac{C_{ox}}{q} \times \Delta V_h, \tag{2}$$

where x indicates the location along the channel, q is the electron charge, L is the channel length, C_{ox} refers to the capacitance of the gate oxide, ΔV_h indicates the differentia of V_h between the test device and its initial state, and N_{ox} is the trapped charge density in the oxide.

Figure 1 exemplifies the CV_b method (testing from source side) with non-uniform localized $V_{T(x)}$ and $V_{FB(x)}$ profiles. Four regions of the channel surface are marked in this figure. V_b is kept lower than V_{FB} and therefore the transistor can be biased in strong accumulation. Before V_h increases to $V_{T(0)}$, which is the lowest V_T along the channel, the channel will not turn inversion and no I_{cp} can be generated. When V_h is set lower than the highest $V_{T(x)}$ but higher than $V_{T(0)}$, taking the V_T of point a ($V_{T(a)}$) indicated by the dashed line as an example, regions A and A′ will become inversion and trap electrons in the interface states. The subsequent discharged electrons in region A can be sensed by the source side; however, the electrons in region A′ are blocked by the wide V_T barrier and thus cannot be collected. When V_h reaches the highest V_T (set as $V_{T(b)}$), the whole channel turns on and thus all the other interface states can contribute to the I_{cp}. Assuming that the interface states are symmetrically distributed, CP current will double and reach saturation. The precise V_h versus $V_{T(x)}$ relationship can be determined by the measured I_{cp}–V_h curve from both drain and source sides.

However, using this CV_b technique to measure the trapped charge distribution on advanced devices with an effective channel length of less than 100 nm may result in unacceptable errors. Figure 1(b) illustrates a narrow V_T/V_{FB} peak in a short channel MOSFET to explain this phenomenon. As the channel length shrinks, the distribution of the localized V_T peak is in the nanometer scale. Before V_h reaches $V_{T(b)}$, the drain side charge pumping current is hard to effectively block by the super narrow $V_{T(x)}$ barrier due to the punch through effect, and join to the $I_{cp,s}$ tested at the source. As illustrated in Fig. 1(b), point a is the location where the punch through effect occurs. This punch through current refers to the segment A′ in the I_{cp}–V_h curve in Fig. 2(b), and thus the punch through point in the channel can be deduced directly from the I_{cp} figure. With the continued growth of V_h, interface states in regions B and B′ lead to the third segment in the I_{cp} curve. As a result, the I_{cp}–V_h curve acquired from the short channel MOSFET which contains a narrow V_T peak is different from the traditional I_{cp} curve. The peak of the V_T and corresponding trapped charge profile is not accurate. On the other hand, the I_{cp} curve shift caused by the

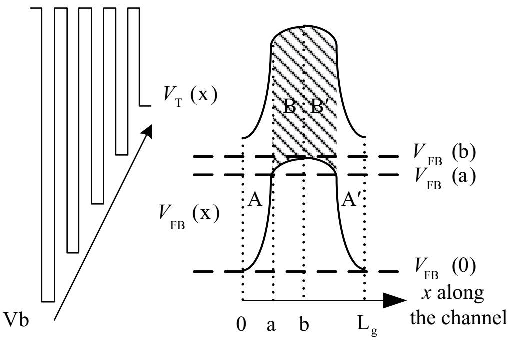

Fig. 3. Illustration of the CV_h method in a short channel MOSFET.

V_T peak is in the large current region, which makes it even less precise. Consequently, the conventional CV_b technique is not appropriate for characterizing the trapped charge profile for nanometer-scale devices.

The so-called "constant V_h (CV_h)" method can also be used to extract the profile of the V_T peak, in which the high level of the gate pulse is fixed while the base value is varied[5]. Figure 3 shows the CV_h method applied in the same short channel MOSFET. V_h is set larger than the highest V_T along the channel. When the gate pulse is held as V_h, the whole channel is pulsed into inversion. When V_b reaches the V_{FB} of point a ($V_{FB(a)}$) which is higher than the V_{FB} of regions A and A′, these two regions cannot turn accumulation and form charge pumping current, while regions B and B′ which contain a V_{FB} peak which is higher than V_b will generate charge pumping current. The narrow V_T/V_{FB} peak can be measured and does not affect the precision of CP testing. Assuming that the surface states are uniformly distributed, the measured charge pumping current is directly proportional to the width of the V_{FB} peak. The corresponding relationship of V_{FB} and charge pump current in CV_h can be used to determine the trapped charge density. By replacing V_h with V_b, Equations (1) and (2) can be used in CV_h to get the $N_{ox,\,peak}$ curve:

$$w = L\frac{I_{cp}(V_b)}{I_{cp,\,max}}, \tag{3}$$

$$N_{ox,\,peak}(w) = \frac{C_{ox}}{q} \times \Delta V_b, \tag{4}$$

where w indicates the V_{FB} peak width, ΔV_b indicates the differentia of V_b between the test device and its intial state, and $N_{ox,\,peak}$ is the peak density of the trapped charge in the oxide.

The curve shift caused by the V_{FB} peak takes place in the low current region and has higher precision in the CV_h method. Because V_h is set larger than the highest V_T along the whole channel, I_{cp} current cannot correspond to a specific x location along the channel. Therefore, CV_h can only be used to extract the width and value of the trapped charge peak, but cannot locate the position of the charge peak and further profile of the trapped charge distribution.

In this paper, a novel modified CP method is proposed, which combines both CV_b and CV_h to obtain the more precise distribution of nanometer-scale devices. The trapped charge profile in regions A and A′ (which refers to the lower V_T and V_{FB}) can be extracted accurately by the CV_b method from the source/drain side, while the width and density value of trapped charges in regions B and B′ (which refers to the narrower V_T and V_{FB} peak) can be extracted by the CV_h method. As a result, the combination of the two methods can be used to characterize the position and density of trapped charges.

3. Results and discussion

Nonvolatile devices based on storage in nitride traps have a localized charge trapping characteristic. On the other hand, the width and density of the trapped charges can be simply controlled by transient program operation. As a result, a SONOS device is introduced to verify the proposed CP method. The sample's W/L is 1 μm/0.12 μm, which is operated by the transient FN program and BBHH erases from the drain and source sides that will not increase the interface states substantially. During the erase period, hot holes injected from the drain and source overlap into the nitride layer and pull down the localized V_T. The rest of the V_T peak is narrower than 100 nm and even 50 nm as the erase time increases. The chosen 0.12 μm device obtains a narrow V_T peak whose width is controllable by erase operation and thus is appropriate to be used to examine the conventional CV_b and the proposed novel method. In addition, a 1 μm/1 μm SONOS is also measured as the control experiment.

First, I_{cp} was measured in the virgin cell. Then, both cells were programmed by FN tunneling injection (V_g = 15 V) for 5 ms and then erased by band-to-band hot hole injection (V_g = –7 V, V_d = 3.6 V and V_g = –7 V, V_s = 3.6 V) from both source side and drain side for 100 μs, 1 ms and 10 ms respectively. I_{cp} curves were measured after each program and erase operation. The CP measurements were set as follows: V_b is fixed at –6 V while V_h sweeps from –4 to 4 V in the CV_b method, and V_h is fixed at 4 V while V_b sweeps from –2 to 3 V in the CV_h method. $I_{cp,\,d}$ is measured with V_s floating and V_d = 0, and vice versa with $I_{cp,\,s}$. The frequency of the gate pulse is set to 5 MHz in order to gain a relative large charge pumping current which may improve the measurement accuracy.

The measured I_{cp}–V_h curves of SONOS cells after the FN program and BBHH erase are shown in Fig. 4. It is observed that the actual I_{cp} curves for both long channel and short channel SONOS are identical to the theoretically analyzed illustration shown in Fig. 2. Injected holes near the source/drain result in lower localized V_T and V_{FB}, which in turn lead to the leftward shift of the I_{cp} curves in the low current region. For the 1 μm SONOS, the middle of the channel is not affected by narrow trapped holes and thus the high current region of I_{cp} remains the same. However, for the 0.12 μm SONOS, trapped holes are distributed widely relative to the narrow channel and ensure that the entire channel is erased thoroughly. The punch through effect can be clearly observed which approximately doubles I_{cp} due to the symmetrical V_T profile in the source and drain sides. In addition, Figure 5 gives the experimental results of the 0.12 μm SONOS in the CV_h method.

The trapped charge density in the nitride layer (N_{ntr}) distribution along the channel direction can be obtained by the CV_b method in Figs. 6(a) and 6(b) for the 1 μm and 0.12 μm SONOS. The zero reference level refers to the virgin state; the net negative N_{ntr} means trapped holes and the net positive N_{ntr} refers to trapped electrons. Two bundles of holes are trapped near the drain and source regions into the nitride layer.

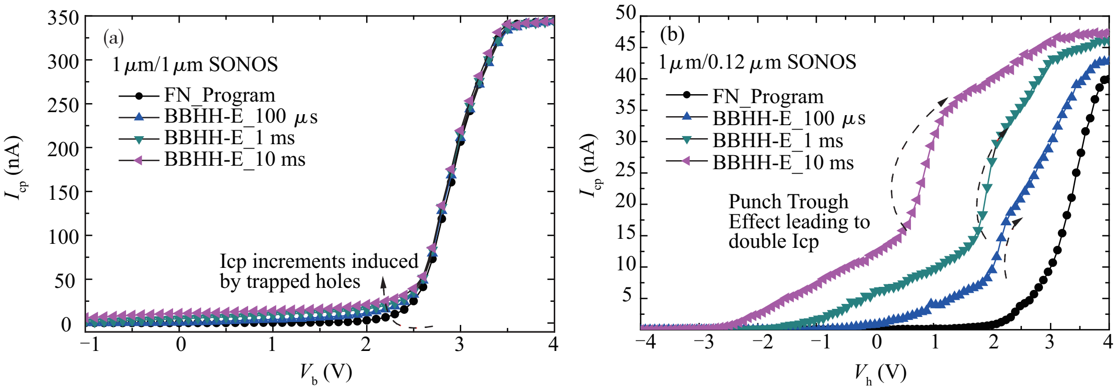

Fig. 4. I_{cp}–V_h curves of (a) 1 μm and (b) 0.12 μm SONOS cell after FN program and BBHH erase.

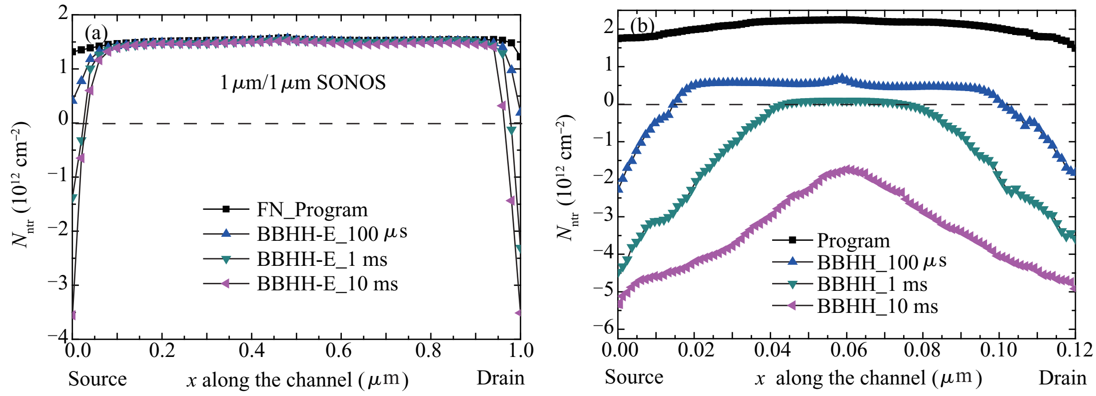

Fig. 6. N_{ntr} versus x curves of (a) 1 μm and (b) 0.12 μm SONOS cell calculated from I_{cp} by the CV$_b$ method.

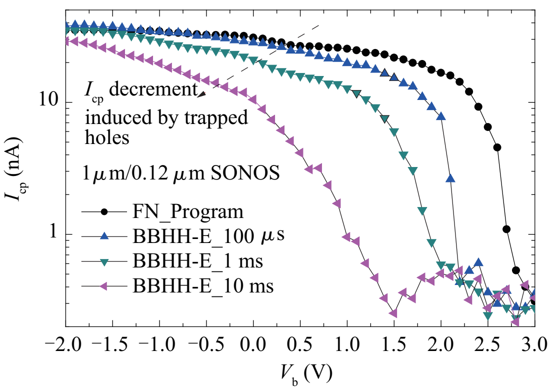

Fig. 5. I_{cp}–V_b curves of 0.12 μm SONOS cell after FN program and BBHH erase.

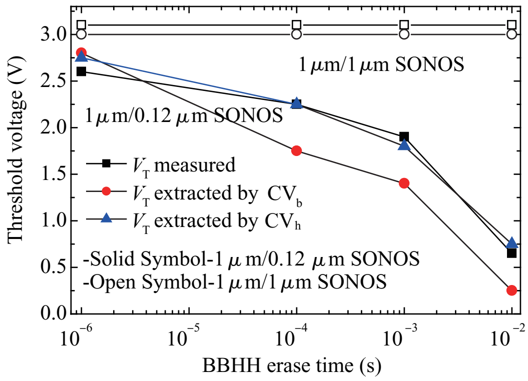

Fig. 7. V_T of 1 μm (open symbols) and 0.12 μm (solid symbols) SONOS cells which are measured and extracted from the I_{cp} curves by the CV$_b$ and CV$_h$ methods.

The width of trapped holes enlarges with extended erase time, therefore the rest of the V_T peak of the 0.12 μm SONOS is very narrow. For the 1 μm SONOS, the extracted trapped holes show a nearly Gaussian distribution. However, the extracted trapped holes in the 0.12 μm SONOS are far from being Gaussian distributed in the peak area, which seems questionable. In addition, Figure 7 shows the comparison of V_T measured by the traditional technique and that extracted from I_{cp} curves by the CV$_b$ and CV$_h$ methods. V_T extracted by CV$_b$ is similar to the measured V_T concerning the long channel device. It is worth noting that V_T extracted by CV$_h$ is more precious than V_T extracted by CV$_b$ for the short channel SONOS. The peak of the trapped holes, which reveals the highest V_T along the whole channel, is obviously pulled down due to the I_{cp} punch through effect arising in the CV$_b$ method. Consequently, although a satisfactory result is obtained using the CV$_b$ method for the long channel device, it encounters problems in the short channel device and should be modified by the CV$_h$ method.

The dashed line in Fig. 8(a) shows the narrow peak profile deduced from the I_{cp}–V_b curve. In contrast, CV$_h$ can extract the width and value of the charge peak, but not the ac-

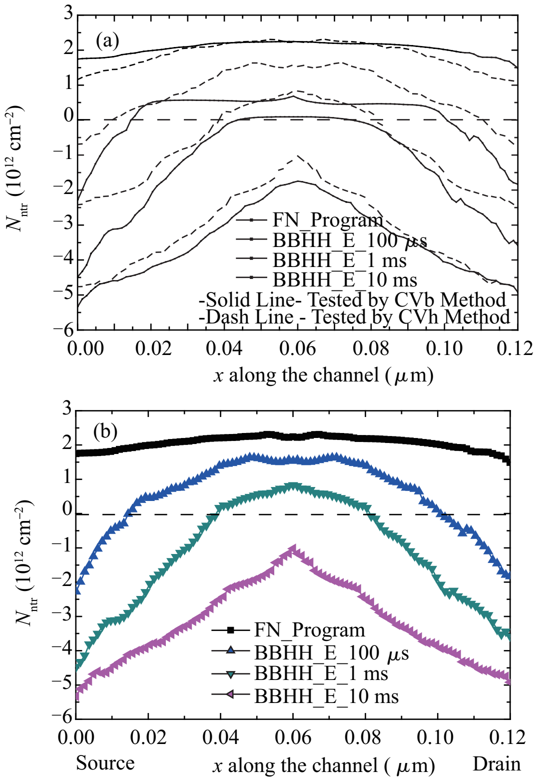

Fig. 8. $N_{\rm ntr}$ versus x of the 0.12 μm SONOS cell calculated from $I_{\rm cp}$ by $\rm CV_b$ (solid line) and $\rm CV_h$ (dash line). (b) Trapped charge distribution along the channel direction in the 0.12 μm SONOS cell.

curate location. In Fig. 8(a), curves obtained from $\rm CV_b$ and $\rm CV_h$ intersect half way up the charge peak, above which the $\rm CV_b$ method cannot acquire an accurate distribution. The narrow charge peak obtained with the $\rm CV_h$ method is inserted into the punch through area to form the whole profile of the trapped charges. As a result, the combination of the two methods can be utilized to extract the whole profile of the trapped charge in the deep submicron transistor.

The final result of the distribution of injected charges along the channel is shown in Fig. 8(b). The acquired trapped charge distribution show good consistency with the measured $V_{\rm T}$ in the $I_{\rm d}$–$V_{\rm g}$ curves. The distribution of the trapped charges may bring out further understanding about the transient characteristics of BBHH injection, which helps to improve the program/erase cycling reliability[10] and 2-bit characteristics[11] of the charge trapping device.

4. Conclusion

In this paper, a new "combination" method based on the CP technique, which combines both the $\rm CV_b$ and $\rm CV_h$ methods, is proposed to study the distribution profile of injected charges in nanometer-scale devices. Two SONOS devices with long and short channel lengths are measured with the traditional and novel methods. Accurate trapped electrons and hole profiles during programming and erase are obtained by the proposed method. This modified method can help to further understand and model the degradation characteristics of an advanced MOSFET caused by hot charges and the charge trapping characteristics of charge trapping memory.

References

[1] Heremans P, Maes H E, Saks N. Evaluation of hot carrier degradation of n-channel MOSFET's with the charge pumping technique. IEEE Electron Device Lett, 1986, EDL-7: 428
[2] Fair R B, Sun R C. Threshold-voltage instability in MOSFET's due to channel hot-hole emission. IEEE Trans Electron Devices, 1981, ED-28: 83
[3] Hofmann K R, Werner C, Weber W, et al. Hot-electron and hole-emission effects in short n-channel MOSFET's. IEEE Trans Electron Devices, 1985, ED-32: 691
[4] Tsuchiaki M, Hara H, Morimoto T, et al. A new charge pumping method for determining the spatial distribution of hot-carrier-induced fixed charge in p-MOSFETs. IEEE Trans Electron Devices, 1993, 40: 1768
[5] Groeseneken G, Maes H E, Beltran N, et al. A reliable approach to charge-pumping Measurements in MOS Transistors. IEEE Trans Electron Devices, 1984, ED-31: 42
[6] Vuillaume D, Marchetaux J C, Boudou A. Evidence of acceptor-like oxide defects created by hot-carrier injection in n-MOSFETs a charge-pumping study. IEEE Trans Electron Devices, 1991, 12: 60
[7] Chen W, Balasinski A, Ma T P. Lateral profiling of oxide charge and interface traps near MOSFET junctions. IEEE Trans Electron Devices, 1993, 40: 187
[8] Chen C, Ma T P. Direct lateral profiling of both interface traps and oxide charge in thin gate MOSFET devices. IEEE VLSI Technol Symp, 1996: 230
[9] Gaitan M, Enlow E W, Russell T J. Accuracy of the charge pumping technique for small geometry MOSFETs. IEEE Trans Nucl Sci, 1989, 36: 1990
[10] Mahapatra S, Shukuri S, Bude J. CHISEL flash EEPROM. II. reliability. IEEE Trans Electron Devices, 2002, 49: 1302
[11] Wang T, Tsai W J, Gu S H, et al. Reliability models of data retention and read-disturb in 2-bit nitride storage flash memory cells. IEDM Tech Dig, 2003: 7.4.1