

Study of the Gate-Sensing and Channel-Sensing Transient Analysis Method for Monitoring the Charge Vertical Location of SONOS-Type Devices

Pei-Ying Du, Hang-Ting Lue, *Member, IEEE*, Szu-Yu Wang, Erh-Kun Lai, Tiao-Yuan Huang, *Fellow, IEEE*, Kuang-Yeu Hsieh, Rich Liu, *Senior Member, IEEE*, and Chih-Yuan Lu, *Fellow, IEEE*

Abstract—The gate-sensing and channel-sensing transient analysis method is studied in detail. This method introduces an additional gate-sensing capacitor to be compared with the conventional channel-sensing one. Sensing in both modes provides two equations that are suitable to solve for two variables—the charge density (Q) and the average charge vertical location (x). In this paper, the principle of this method is discussed in detail. Several factors that affect the measurement accuracy are also analyzed. The power of this method is demonstrated by program/erase cycling and data retention tests. This method is indeed a powerful tool for detailed understanding of trapping dynamics.

Index Terms—Charge density, silicon-oxide-nitride-oxide-silicon (SONOS), transient analysis, trapping dynamics, vertical location.

I. INTRODUCTION

THE future NAND Flash memory is forecasted to adopt the silicon-oxide-nitride-oxide-silicon (SONOS)-type devices in order to solve the scaling issues of floating-gate devices [1]. However, the understanding of nitride traps is very limited so far. In order to make SONOS-type NAND Flash more feasible, the understanding of nitride trap is necessary.

The first thing to understand nitride trap is to find out the trapped charge vertical location. Suzuki *et al.* [2] used metal-oxide-nitride-oxide-silicon (MONOS) capacitors and compared the saturated flat-band voltage shift under FN programming. In their method, they assume that the trap density remains the same for various ONO thicknesses during theoretical calculation. By comparing the theoretical calculation to experimental results, they observed that most of the measured traps are the interface traps between top oxide and nitride. Ishida *et al.* [3] also used a similar method to compare various ONO capacitors. In this method, substrate avalanche hot-electron injection

(which is not a popular operation method for memory devices) is applied to program SONOS capacitors to saturated V_{FB} . By comparison of various thickness, they also concluded that the electrons are mostly trapped at the interface of nitride and oxide, whereas bulk electron traps are negligible. In the two previously mentioned methods, the saturated trap density of various ONO capacitors is assumed to be the same. However, the saturated V_{FB} of a capacitor is also a function of applied voltage; thus, these results may be equivocal. Moreover, these methods do not provide the time evolution of charge vertical location during program/erase or retention testing and only give an estimate of the saturated trap density.

Previously, we proposed a transient analysis method [4] to examine the trap vertical location. We transformed the Vt -time curve (during gate injection) into the transient current (J) and instantaneous electric field (E). By comparing the $J-E$ curves for various possible charge locations, the case with consistent $J-E$ plots then indicates the true charge vertical location. The advantage of this method is that it uses only a single ONO condition without comparing various samples, thus avoiding ambiguity. However, the injected charge must be assumed to be fully captured, which may not be the correct assumption for some samples (for example, a thin nitride). Moreover, the time evolution of charge still cannot be extracted, and this method only provides a rough estimate of the charge location.

The limit of all the aforementioned techniques originates from the difficulty of finding both the charge (Q) and the vertical location (x) using only one equation. Another equation must be added to solve for both variables.

A simple method is to directly measure the injection current during direct-tunneling (DT) programming for a capacitor [5]. The time integration of the injection current gives the total charge Q . By solving Q and V_{FB} shift, the charge vertical location (x) can be found. However, this method also assumes that the injection current is fully captured by SONOS devices. Moreover, the injection current during DT programming is a fast transient response ($\sim \mu s$), which is difficult to measure accurately, and the extracted x is very sensitive to the errors on the detected Q . An accurate measurement requires a complicated setup to compensate for noise. Moreover, this method cannot be applied to the long-term retention test because the de-trapping current is too small for detecting. These limitations greatly affect the accuracy of this method.

We proposed a much simpler way to solve the aforementioned difficulties [6]. We introduced an additional

Manuscript received July 9, 2007.

P.-Y. Du is with the Emerging Central Laboratory, Macronix International Company, Ltd., Hsinchu 300, Taiwan, R.O.C., and also with the Department of Electronics Engineering, National Chiao-Tung University, Hsinchu 300, Taiwan, R.O.C. (e-mail: pennydu@mxic.com.tw).

H.-T. Lue, E.-K. Lai, K.-Y. Hsieh, R. Liu, and C.-Y. Lu are with the Emerging Central Laboratory, Macronix International Company, Ltd., Hsinchu 300, Taiwan, R.O.C.

S.-Y. Wang is with the Technology Development Center, Macronix International Company, Ltd., Hsinchu 300, Taiwan, R.O.C.

T.-Y. Huang is with the Department of Electronics Engineering, National Chiao-Tung University, Hsinchu 300, Taiwan, R.O.C.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TDMR.2007.907290

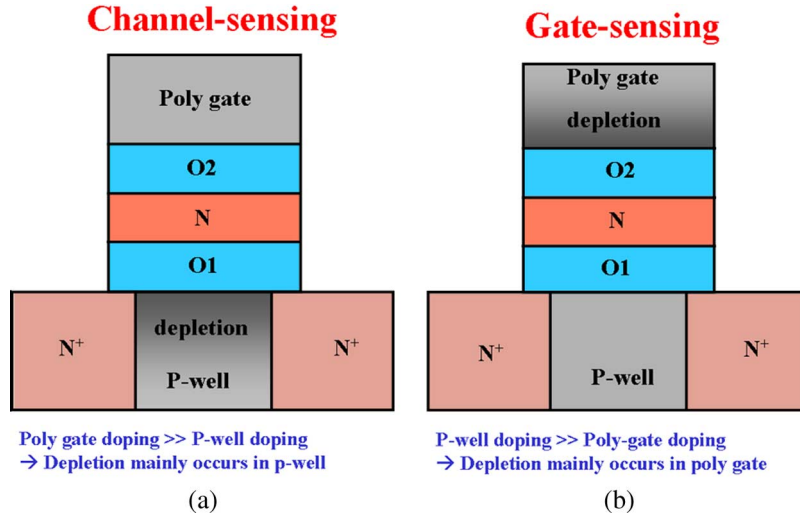


Fig. 1. Comparison of the (a) channel-sensing and (b) gate-sensing methods.

TABLE I
SAMPLE DESCRIPTIONS. FOR EACH SAMPLE, BOTH GATE-SENSING (“pl”) AND CHANNEL-SENSING (“ch”) CAPACITORS ARE MADE

		P-Well (cm ⁻³)	P-poly gate (cm ⁻³)	O1 (Å)	N (Å)	O2 (Å)
S1(SONOS)	pl	7E18	5E16	54	70	90
	ch	1E17	2E20	54	70	90
S2(BE-SONOS)	pl	7E18	5E16	15/20/25	70	90
	ch	1E17	2E20	15/20/25	70	90

gate-sensing capacitor to be compared with the conventional channel-sensing one. The sensing in both modes ($V_{FB,ch}$ and $V_{FB,pl}$) provides two equations to solve Q and x simultaneously. Only simple capacitance–voltage (C – V) measurement is required, and there is no complicated setup or noise compensation. Moreover, we also provide a “real-time” measurement method since Q and x are tracked during program/erase and retention testing. In this paper, the principle of this method is discussed in detail, and several factors that affect the measurement accuracy are examined. The power of this method is demonstrated by program/erase cycling and data retention tests.

II. METHOD AND THEORY

The proposed method measures two ONO capacitors: one using channel-sensing and the other using gate-sensing, as shown in Fig. 1. The two capacitors have identical ONO but different doping conditions. The channel-sensing capacitor has a lightly doped p-well and a heavily doped p⁺-poly gate. Therefore, the depletion mainly occurs in the p-well, and the flat-band voltage shift ($\Delta V_{FB,ch}$) is sensed by the channel depletion. On the other hand, the gate-sensing capacitor has a lightly doped p-poly gate and a heavily doped p-well. Therefore, the depletion mainly happens in the p-poly gate, and the $\Delta V_{FB,pl}$ is sensed by the poly depletion. The typical doping conditions for gate-sensing and channel-sensing capacitors are shown in Table I.

For the same charge configuration inside the nitride, the two capacitors give different V_{FB} shifts and, thus, can tell us Q and x simultaneously. This is because we have two equations (one from channel-sensing and the other from gate-sensing) to solve the two variables. On the other hand, if we only have

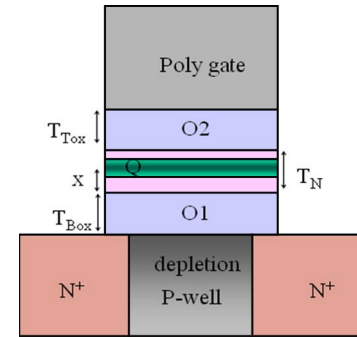


Fig. 2. Definition of the sheet charge density and the relative vertical location.

the channel-sensing capacitor, Q and x cannot be determined simultaneously.

First, we assume a sheet charge inside the nitride for simplicity. The density of the sheet charge is defined as Q (electron per square centimeter), whereas the relative vertical location in nitride is defined as \hat{x} (centimeter) which is measured from N/O1 interface, as shown in Fig. 2. For the channel-sensing method, the $\Delta V_{FB,ch}$ is proportional to the charge density multiplied by the distance toward the polygate [7]

$$\Delta V_{FB,ch} = Q \left(\frac{T_{Tox}}{\epsilon_0 \epsilon_{ox}} + \frac{T_N - \hat{x}}{\epsilon_0 \epsilon_N} \right). \quad (1)$$

On the other hand, for the gate-sensing method, since the channel and gate are effectively reversed, the equation becomes

$$\Delta V_{FB,pl} = Q \left(\frac{T_{Box}}{\epsilon_0 \epsilon_{ox}} + \frac{\hat{x}}{\epsilon_0 \epsilon_N} \right). \quad (2)$$

T_{Tox} is the top oxide (O2) thickness, T_N is the nitride (N) thickness, and T_{Box} is the bottom oxide (O1) thickness, as defined in Fig. 2. ε_0 is the permittivity in vacuum; ε_{ox} and ε_N are the dielectric constants of oxide and nitride, respectively. It should be noted that in (1) and (2), Q is positive for electron charge, whereas it is negative for hole charge.

The V_{FB} shifts can be transformed into Q and \hat{x} by solving (1) and (2)

$$Q = \varepsilon_0 \varepsilon_{\text{ox}} \frac{\Delta V_{\text{FB,ch}} + \Delta V_{\text{FB,pl}}}{\text{EOT}} \quad (3)$$

$$\hat{x} = \frac{\Delta V_{\text{FB,pl}}(T_{\text{Tox}} \cdot \varepsilon_N + T_N \cdot \varepsilon_{\text{ox}}) - T_{\text{Box}} \cdot \Delta V_{\text{FB,ch}} \cdot \varepsilon_N}{\varepsilon_{\text{ox}}(\Delta V_{\text{FB,ch}} + \Delta V_{\text{FB,pl}})}. \quad (4)$$

EOT is the effective oxide thickness of ONO.

Although (3) and (4) are derived by assuming a sheet charge density, the equations are also valid for vertically distributed charges of arbitrary shape, where

$$Q = \int_0^{T_N} Q(x) dx \quad (5)$$

$$\hat{x} = \frac{1}{Q} \int_0^{T_N} Q(x) x dx. \quad (6)$$

Q is the summation of charges throughout the entire trapping layer, and \hat{x} is the mean vertical location inside the nitride. The reason why (3) and (4) can be transformed to the mean value results is simply due to the principle of linear superposition. A detailed derivation is given in Appendix A.

In this method, the well and poly gate doping does not need to be p-type. N-type well and n-type poly gate can be used as well. We also used wells and gates with opposite polarities. For example, a lightly doped p-poly gate with a heavily doped n-well was used for the gate-sensing capacitor. However, the measured C - V curve had a larger distortion (data are not shown here), which affected the accuracy of V_{FB} shift. This is because both poly depletion and substrate depletion happened at the same time during $-V_{\text{gate}}$ bias for this device, resulting in a complicated C - V curve. Therefore, it is best to use the same polarity for both well and poly gate because only one of them has depletion during $\pm V_{\text{gate}}$ bias.

III. SAMPLE DESCRIPTION

Large area ($500 \mu\text{m} \times 500 \mu\text{m}$) SONOS capacitors with various ONO conditions are fabricated. For each sample, both gate-sensing (“pl”) and channel-sensing (“ch”) capacitors are made. Table I illustrates the ONO thickness and the estimated doping concentration for each sample. S2 is a bandgap-engineered SONOS (BE-SONOS) device [8], where O1 is replaced by a multi-layer ONO barrier ($15/20/25 \text{ \AA}$). For all devices, n^+ -doped source/drain regions are formed so that channel inversion layer can be formed during +FN programming.

The bottom oxide (O1) is formed by wet oxidation, whereas the nitride layer is deposited by low-pressure chemical vapor

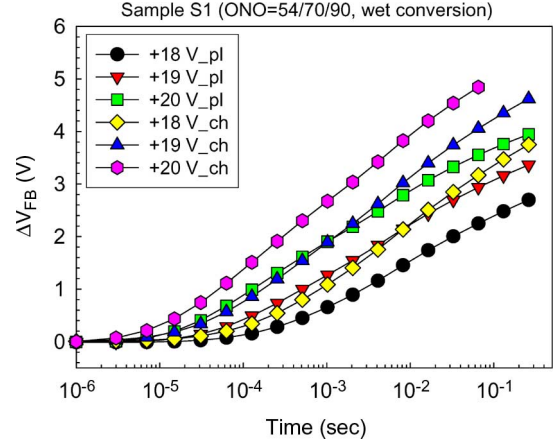


Fig. 3. Comparison of the V_{FB} shifts during +FN programming. “pl” indicates the gate-sensing results, whereas “ch” indicates the channel-sensing results.

deposition (LPCVD) process using dichlorosilane (DCS) and NH_3 . For the top oxide (O2), high-temperature ($1000 \text{ }^\circ\text{C}$) wet oxidation of nitride (wet conversion) is used.

The ONO of the channel-sensing and gate-sensing capacitors are grown simultaneously in order to provide identical thickness.

IV. EXPERIMENTAL RESULTS

Sample S1 is programmed by +FN using various bias voltages from the initial fresh state ($V_{\text{FB}} \sim 0 \text{ V}$), and both the channel-sensing and gate-sensing capacitors are measured under identical electrical testing procedures. The results of V_{FB} shifts and C - V curves are shown in Figs. 3 and 4, respectively. In Fig. 3, the V_{FB} shifts of gate-sensing and channel-sensing are clearly different. This is because $\text{O2} > \text{O1}$ for sample S1. According to (1) and (2), $\Delta V_{\text{FB,ch}}$ should be larger than $\Delta V_{\text{FB,pl}}$.

The corresponding C - V curves are shown in Fig. 4. In Fig. 4(a), the C - V curves of channel-sensing capacitor are in accumulation at $-V_{\text{gate}}$ (or $+V_{\text{well}}$) and become depletion and even inversion at $+V_{\text{gate}}$ (or $-V_{\text{well}}$). The inversion layer can be formed because n^+ source/drain regions are introduced. On the other hand, the C - V curves of gate-sensing capacitor are in accumulation for $-V_{\text{well}}$ (or $+V_{\text{gate}}$) and become depletion at $+V_{\text{well}}$ (or $-V_{\text{gate}}$). The inversion layer cannot be formed because the poly gate does not have source/drain regions; hence, deep depletion continues at larger $+V_{\text{well}}$. Moreover, the C - V curves of gate-sensing capacitor have much more gradual slope during transition. This is possibly because there are more interfacial states at poly gate and top oxide interface. However, the C - V curves still show parallel shift during +FN programming, which is sufficient for V_{FB} shift measurement.

In Fig. 3, the two V_{FB} shifts ($\Delta V_{\text{FB,ch}}$ and $\Delta V_{\text{FB,pl}}$) at a given time can be transformed into Q and x by using (3) and (4); the results are shown in Fig. 5. The charge evolution (Q - t) and location evolution (x - t) can be extracted by this method. In Fig. 5, Q - t has a similar shape as $\Delta V_{\text{FB}} - t$ (Fig. 3), as expected. On the other hand, x - t shows that electrons migrate from the bottom interface toward the top. In order to obtain a

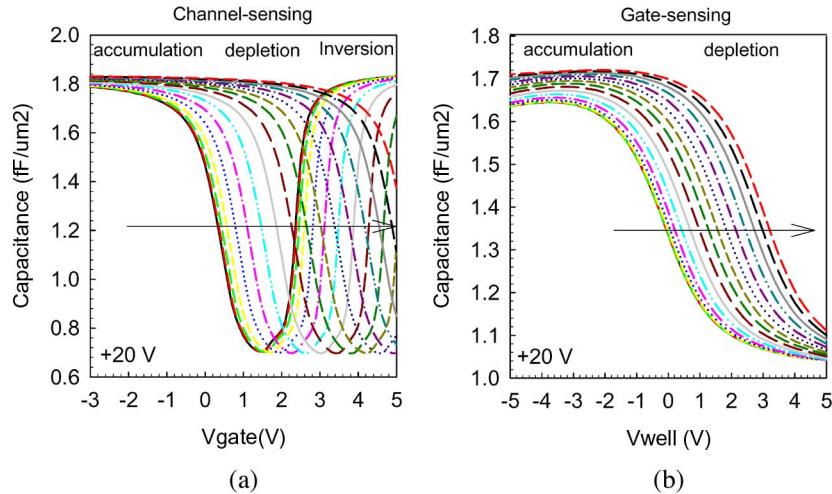


Fig. 4. $C-V$ curves of S1 during +FN programming. (a) Channel-sensing. (b) Gate-sensing.

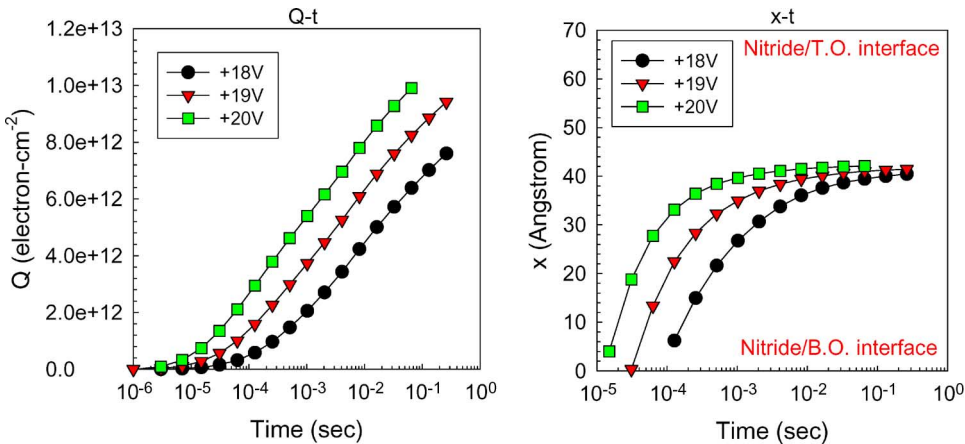


Fig. 5. $Q-t$ and $x-t$ of S1 transformed from Fig. 3 by using (3) and (4).

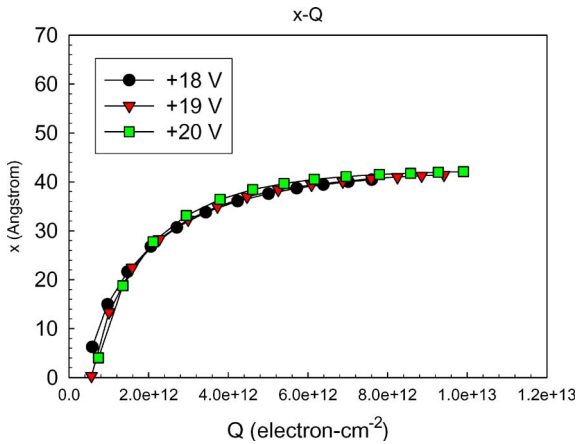


Fig. 6. $x-Q$ of S1 transformed from Fig. 5.

clearer idea about the charge evolution, $x-Q$ curves are plotted, as shown in Fig. 6. $x-Q$ is a simple transformation from $Q-t$ and $x-t$ curves. In Fig. 6, as electron density increases, the mean charge vertical location migrates from the bottom interface toward the center of nitride. The final saturated mean value of vertical location is around 40 Å, which is slightly higher than the center of nitride. Moreover, despite the large dependence of

programming speed to the programming voltage, $x-Q$ plots are independent of the programming voltage. This clearly shows that the $x-Q$ plot expresses an intrinsic property of the ONO structure (the charge trapping behavior) and is not affected by external factors (such as E-field and poly and well doping).

V. COMPARISON WITH OTHER METHODS

In order to verify the charge vertical location in Fig. 6, we use two different methods to monitor the charge location.

The first method is to apply our previous transient analysis method [4], where $J-E$ curves are compared with various charge location assumptions. We plot the injection current density (J) versus bottom oxide electric field (E_{Box}) curves at various programming voltages by assuming different charge vertical locations. The gate-sensing $V_{\text{FB,pl}}$ shifts are used to extract $J-E$ plots, and the detailed equations are shown in Appendix B. For the gate-sensing method, since the channel and gate are effectively reversed, the equations to calculate the top oxide electric field (E_{Tox}) in [4] can be utilized to calculate the E_{Box} in our case. The calculated results are shown in Fig. 7, in which Case 1 represents charges placed at the interface of nitride and top oxide, Case 2 represents charges placed at the interface of bottom oxide and nitride, and Case 3 represents

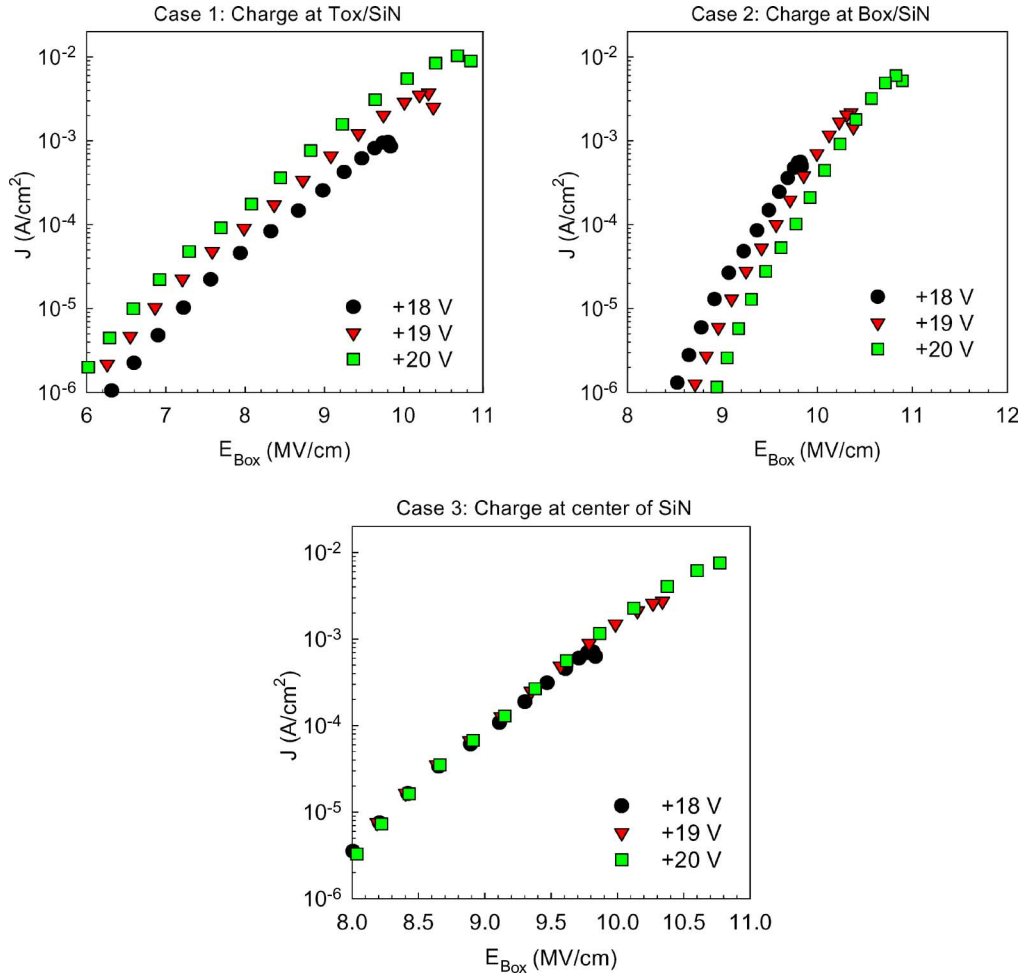


Fig. 7. J versus E_{Box} curves by assuming different charge vertical locations.

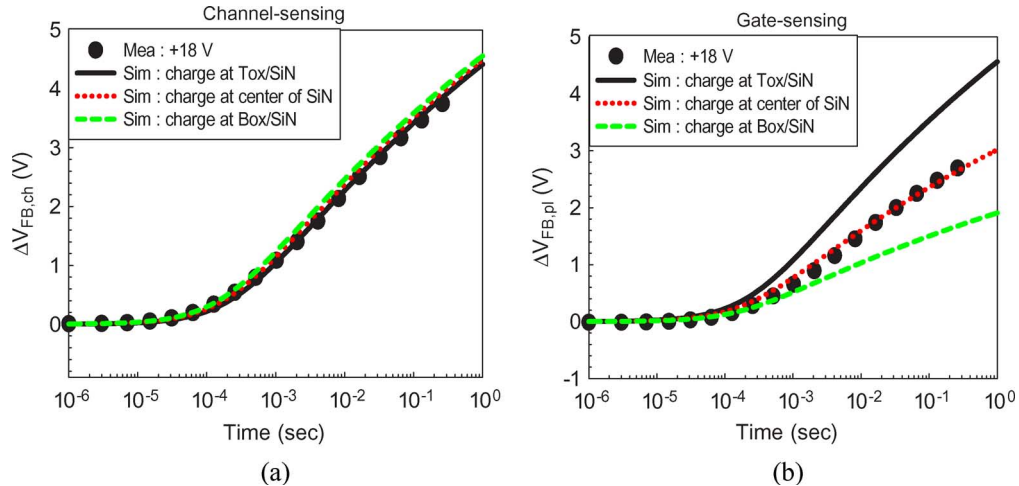


Fig. 8. Comparison of (a) $V_{\text{FB,ch}} - t$ and (b) $V_{\text{FB,pl}} - t$ with the theoretically calculated results by assuming different charge vertical locations.

charges placed at the center of nitride. The results indicate that Case 3 has better consistency for every programming voltage, whereas it is much less consistent for Cases 1 and 2. These results suggest that the centroids of electrons are more likely located near the center of nitride, which is consistent with Fig. 6.

The second method is to directly simulate the $\Delta V_{\text{FB}} - t$ curves assuming various charge vertical locations. The theoretical calculation can be carried out by integrating the FN tunneling current with respect to time. Appendix C illustrates the detail of this simulation method for $\Delta V_{\text{FB}} - t$. The results are shown in Fig. 8.

For channel-sensing capacitor [Fig. 8(a)], it is very interesting that $\Delta V_{\text{FB, ch}} - t$ is almost the same for various charge vertical locations. Equation (A-17) shows the bottom oxide electric field, which is independent of the charge location (for Case 1, 2, or 3). Therefore, we cannot distinguish the real charge location from $\Delta V_{\text{FB, ch}} - t$.

On the other hand, for the gate-sensing capacitor [Fig. 8(b)], since the channel and gate are effectively reversed, the equations of bottom oxide electric field are very different for different locations, as shown in (A-9), (A-11), and (A-13). Fig. 8(b) shows the simulated $\Delta V_{\text{FB, pl}} - t$ for various charge locations. It shows that only Case 3 (assuming that the charges are placed at the center of nitride) is consistent with the measured results. Again, this result supports that the charge centroid is close to the center of nitride after +FN injection.

Although the two previously mentioned methods can obtain results that are consistent with Section IV, these methods have limitations and can only provide a rough estimate of charge vertical location. On the other hand, the new gate-sensing and channel-sensing method in Sections III and IV is more powerful and provides real-time monitoring of traps.

VI. ACCURACY ESTIMATION

There are two major factors that affect the accuracy of this method. The first one is the error in ONO thickness measurement, and the other is that the injected charges of channel- and gate-sensing capacitors are not identical. We first consider the error in the ONO thickness measurement. The typical error in thickness measurement is usually less than 5%. For sample S1, we assume that the thickness in O1, N1, and O2 have $\pm 2\%$ and $\pm 5\%$ variations from the ideal 54, 70, and 90 Å, respectively. We then calculate Q and x from Fig. 3 by using (3) and (4). The corresponding $x-t$ and $Q-t$ plots are shown in Fig. 9. Fig. 9(a) and (b) shows the bottom oxide thickness variation, Fig. 9(c) and (d) shows the nitride thickness variation, and Fig. 9(e) and (f) shows the top oxide thickness variation. The calculated results show that for all kinds of thickness variation, the mean vertical location (x) remains around the center of nitride with the variation of $< 7\%$, and the total charge (Q) is almost unchanged with the variation of $< 1.5\%$. Because the variation of both x and Q is smaller than 10%, the accuracy of our method is good even when there is finite ($< 5\%$) error in ONO thickness measurement.

Next, we consider the other errors caused by the difference between the channel-sensing and gate-sensing capacitors, such as the bottom oxide thickness being not the same for the two capacitors. Different bottom oxide thickness results in different +FN injection currents and then different injected charges. If the injected charges inside the nitride of the two capacitors are different, they result in inadequate determination of Q and x . In order to estimate this error, we assume two cases. The first one is that O1 of the channel-sensing capacitor has $\pm 2\%$ and $\pm 5\%$ variations from 54 Å, whereas the gate-sensing capacitor has exactly 54 Å O1. The second case is that O1 of the gate-sensing capacitor has $\pm 2\%$ and $\pm 5\%$ variations from 54 Å, whereas the channel-sensing capacitor has 54 Å O1. We estimate the error by calculat-

ing theoretical $\Delta V_{\text{FB}} - t$ for both channel-sensing and gate-sensing capacitors, taking into account the thickness variation, as mentioned earlier. In the $\Delta V_{\text{FB}} - t$ calculation, we assume that the $x-Q$ follows a hyperbolic tangent function, which is obtained from fitting the experimental $x-Q$ plot, as shown in Fig. 15. The detailed equations are shown in Appendix D. The calculated $\Delta V_{\text{FB, ch}} - t$ and $\Delta V_{\text{FB, pl}} - t$ are shown in Fig. 10(a) and (b), respectively.

Using (3) and (4), Fig. 10 can be transformed into Q and x , as shown in Fig. 11. Fig. 11(a) and (b) show that channel-sensing capacitor has $\pm 2\%$ and $\pm 5\%$ variations in bottom oxide thickness from the poly-sensing one, and Fig. 11(c) and (d) show that gate-sensing capacitor has $\pm 2\%$ and $\pm 5\%$ variations in bottom oxide thickness from the channel-sensing one. Both $x-t$ plots in Fig. 11(a) and (c) show large deviation from the ideal case at short programming time ($< 1\text{E-}4$ s), whereas they show small deviation at longer programming time. We conjecture that, at short programming time, the amount of injected charges is very small; thus, the x value is sensitive to bottom oxide thickness variation. On the other hand, at longer programming time, the amount of injected charges is large enough to reveal nitride intrinsic characteristic; therefore, the x value is insensitive to bottom oxide thickness variation. The final x variations are smaller than 7.2% and 1% for channel- and gate-sensing capacitors, respectively. For the gate-sensing method, because the channel and gate are effectively reversed during sensing, it is less sensitive to bottom oxide variation. Both $Q-t$ plots in Fig. 11(b) and (d) show small deviation; the variations are smaller than 2.3% and 1% for channel-sensing and gate-sensing capacitors, respectively. Therefore, we show that in our method, the accuracy of location evolution ($x-t$) at long time and charge evolution ($Q-t$) is very high. However, if the amount of injected charges is very small, location evolution ($x-t$) may have larger deviation.

VII. APPLICATION IN P/E CYCLING AND GATE STRESS

A BE-SONOS device (S2) is used to examine the location evolution during P/E cycling. The results of a 100-P/E cycling test are shown in Fig. 12(a). Fig. 12(b) shows the corresponding Q and x . We can see that x is close to the nitride center for the +FN programmed state, whereas it moves higher for the -FN erased state. Moreover, we observe that, at first (P/E < 10), a small amount of ΔV_{FB} instability seems to be caused by charge vertical mismatch [Fig. 12(b)]. Afterward, ΔV_{FB} , Q , and x are stable. The detailed mechanisms still need to be investigated; however, these observations show how powerful our dual-sensing technique is in the revelation of information that was previously not obtainable.

Next, we examine the charge loss during retention test. Sample S1 is first programmed by +FN to inject electrons [Fig. 13(a)]. Then, a negative V_{gate} (-13 or -14 V) is used to expel electrons (detrapping) out of nitride [Fig. 13(b)]. The calculated $x-Q$ plots are shown in Fig. 14. These results suggest that electrons first de-trap from the bottom interface; therefore, the mean value x moves toward the top interface during $-V_{\text{gate}}$ stressing. The detailed mechanism needs further investigation; however, the current paper again demonstrates

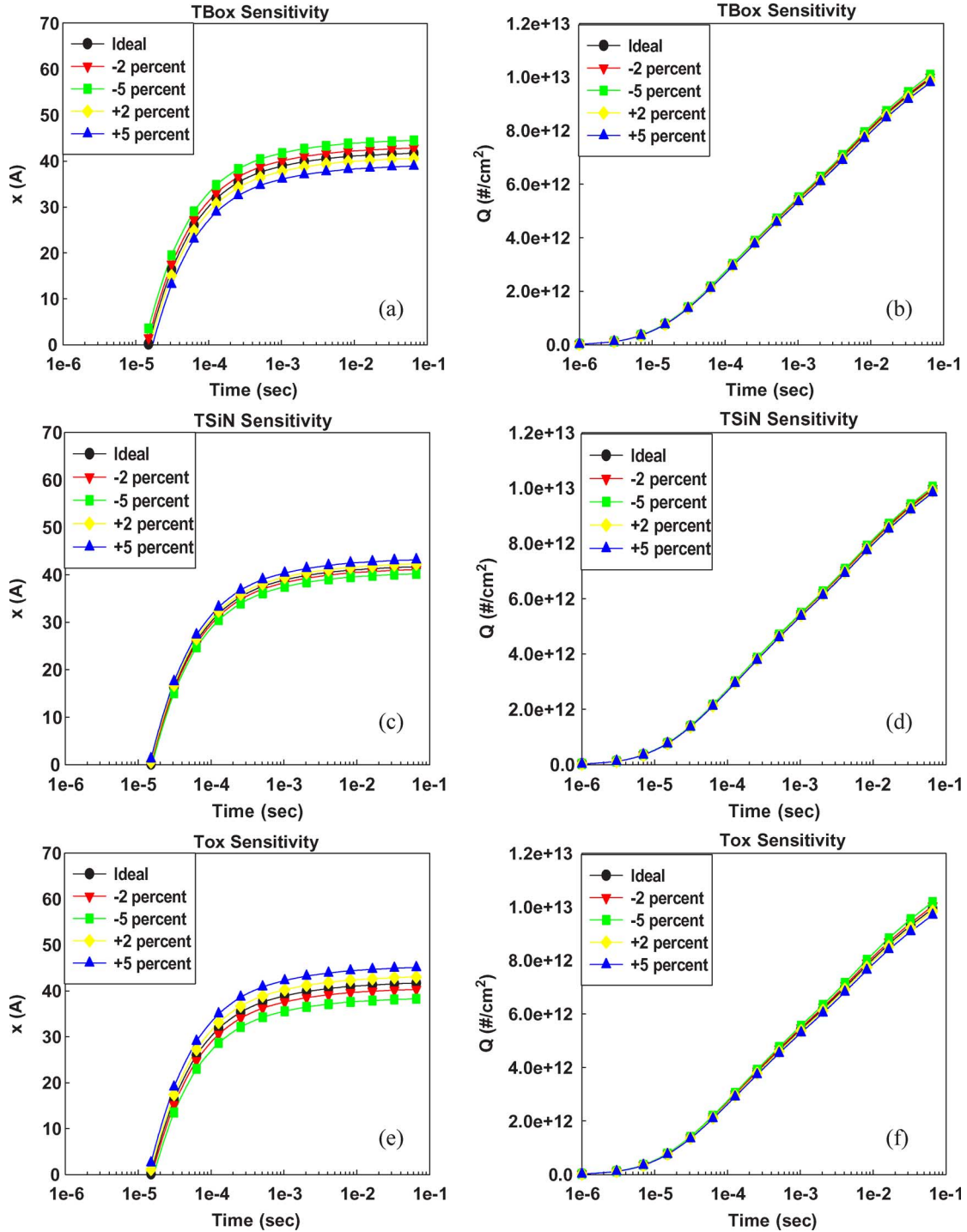


Fig. 9. $x-t$ and $Q-t$ plots for different thickness measurement errors. (a) and (b) show the bottom oxide thickness variation. (c) and (d) show the nitride thickness variation. (e) and (f) show the top oxide thickness variation.

how our new technique helps to obtain new information that was previously unavailable.

VIII. DISCUSSION AND SUMMARY

This paper provides a detailed understanding of the novel gate-sensing and channel-sensing transient analysis method to monitor the trapped charge. By introducing a new measurable quantity ($V_{FB,pl}$), both total injected charge (Q) and mean vertical location (x) can be simultaneously solved during

program/erase or retention test. Thus, this method provides a real-time monitoring of charge injection and migration, which is of crucial help in understanding the detailed physics of traps.

We have extensively examined the accuracy of this new technique. We find that this method has a high accuracy against the thickness error of O1, N1, or O2 thickness. On the other hand, if gate-sensing and channel-sensing capacitors have finite thickness differences in bottom oxide, it shows a larger error in determining the x at initial programming. On the other hand, Q is much less vulnerable to thickness variation.

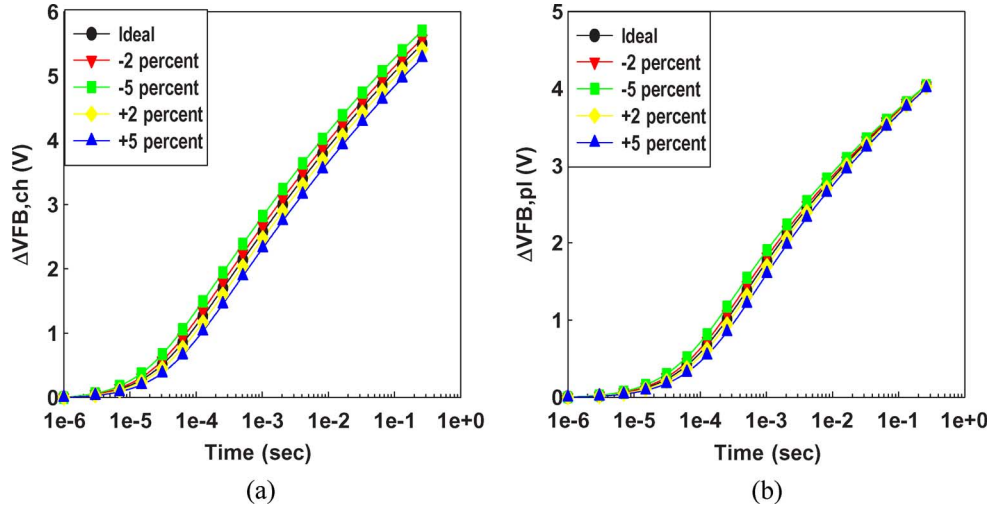


Fig. 10. (a) $V_{FB,ch} - t$ and (b) $V_{FB,pl} - t$ for $\pm 2\%$ and $\pm 5\%$ variations in bottom oxide thickness during +FN programming.

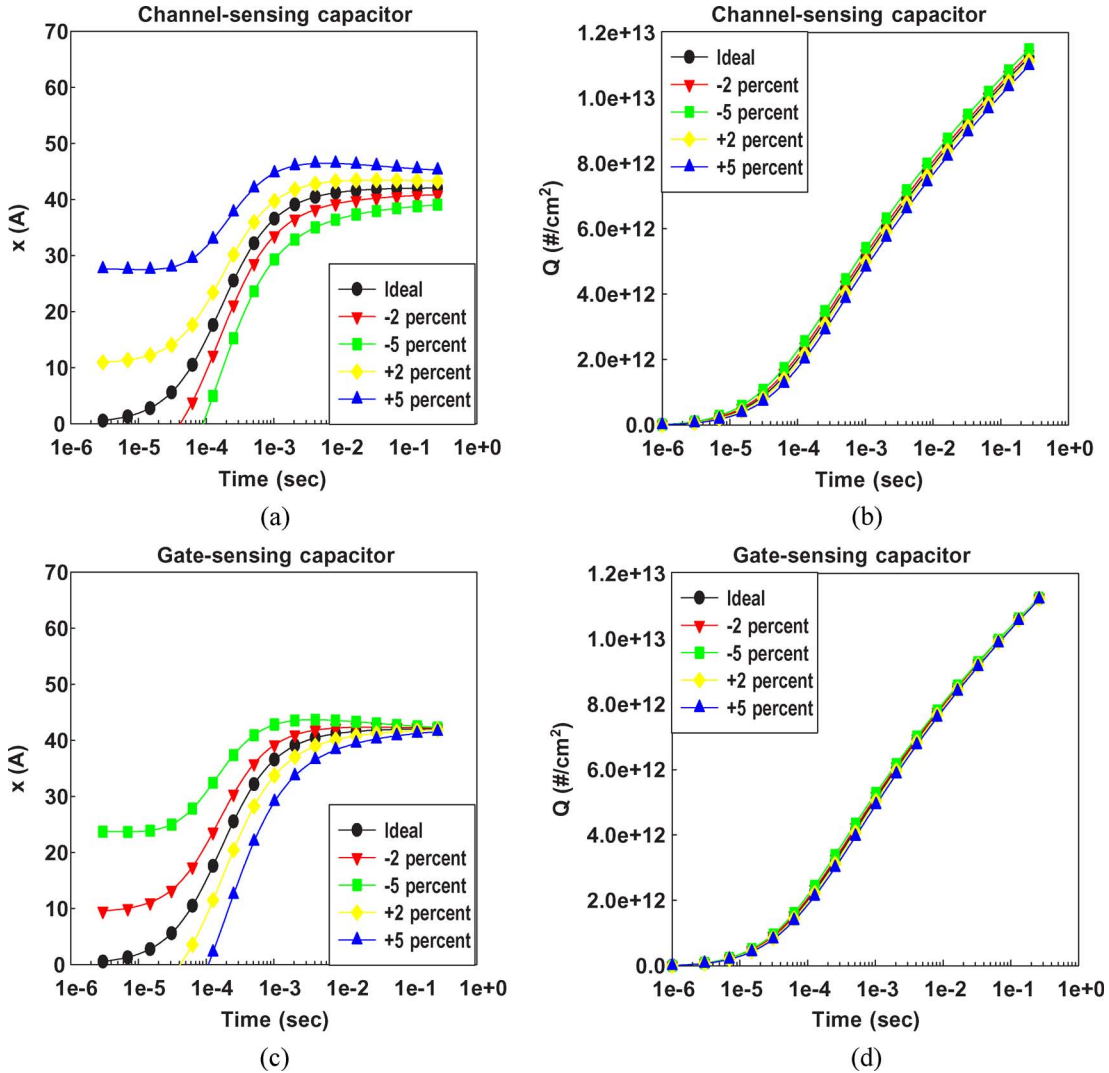
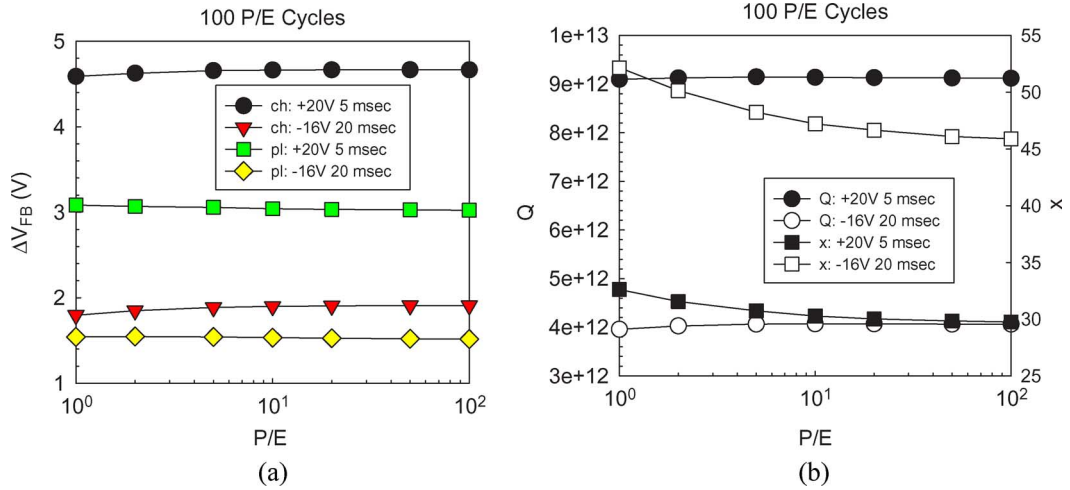
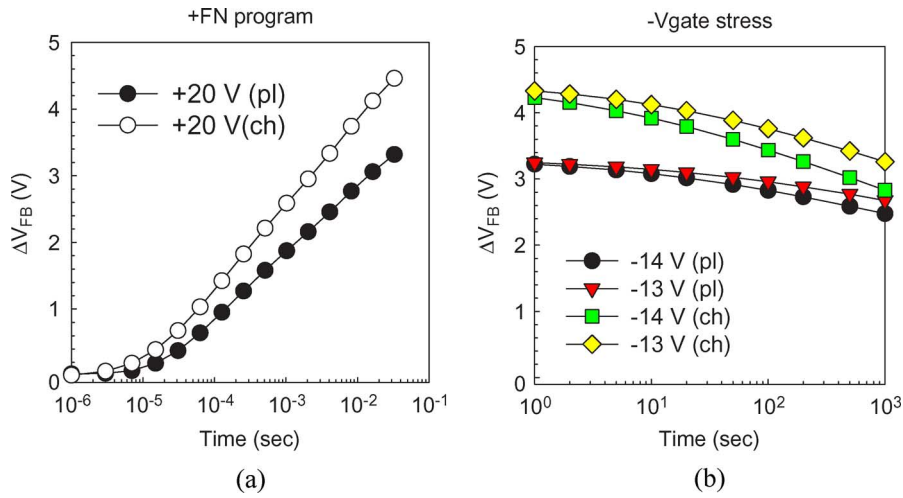


Fig. 11. $x-t$ and $Q-t$ plots transformed from Fig. 10(a) and (b) show that channel-sensing capacitor has $\pm 2\%$ and $\pm 5\%$ variations in bottom oxide thickness from 54 Å. (c) and (d) show that gate-sensing capacitor has $\pm 2\%$ and $\pm 5\%$ variations in bottom oxide thickness from 54 Å.

We also provide several examples of reliability testing of SONOS-type devices. We find that this method provides a detailed examination of charge vertical location and migration

during program/erase or reliability testing, which are hardly measured before. Therefore, this new technique is very powerful in discovering the physics of traps.


 Fig. 12. (a) P/E cycling test of S2 for both channel-sensing and gate-sensing methods. (b) Extracted Q and x during P/E endurance.

 Fig. 13. V_{FB} shifts of (a) programming by +20 V and (b) retention under $-V_{gate}$ stress.

APPENDIX A

DETAILED DERIVATION OF (3) AND (4)

If the charge vertical distribution function is $Q(x)$, then the flat-band voltage shift can be written as

$$\Delta V_{FB,ch} = \int_0^{T_N} Q(x) \left(\frac{T_{Tox}}{\epsilon_0 \epsilon_{ox}} + \frac{T_N - x}{\epsilon_0 \epsilon_N} \right) dx \quad (A-1)$$

$$\Delta V_{FB,pl} = \int_0^{T_N} Q(x) \left(\frac{T_{Box}}{\epsilon_0 \epsilon_{ox}} + \frac{x}{\epsilon_0 \epsilon_N} \right) dx. \quad (A-2)$$

Define the mean value \hat{x} and total charge Q

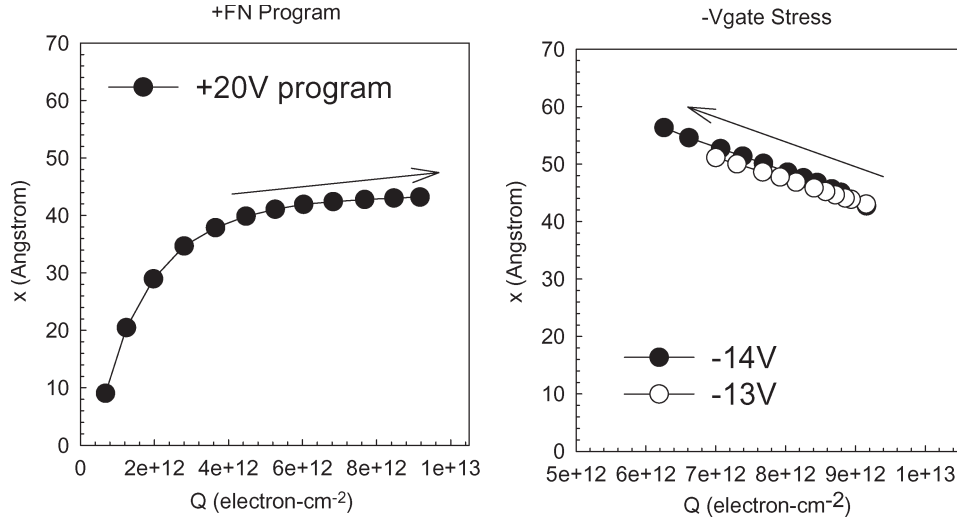
$$Q = \int_0^{T_N} Q(x) dx \quad (A-3)$$

$$\hat{x} = \frac{1}{Q} \int_0^{T_N} Q(x) x dx \quad (A-4)$$

$$\begin{aligned} \Rightarrow \Delta V_{FB,ch} + \Delta V_{FB,pl} &= \int_0^{T_N} Q(x) \left(\frac{T_{Tox} + T_{Box}}{\epsilon_0 \epsilon_{ox}} + \frac{T_N}{\epsilon_0 \epsilon_N} \right) dx \\ &= Q \frac{EOT}{\epsilon_0 \epsilon_{ox}} \end{aligned} \quad (A-5)$$

$$\begin{aligned} \Rightarrow \Delta V_{FB,ch} &= \int_0^{T_N} Q(x) \left(\frac{T_{Tox}}{\epsilon_0 \epsilon_{ox}} + \frac{T_N - x}{\epsilon_0 \epsilon_N} \right) dx \\ &= Q \left(\frac{T_{Tox}}{\epsilon_0 \epsilon_{ox}} + \frac{T_N - \hat{x}}{\epsilon_0 \epsilon_N} \right) \end{aligned} \quad (A-6)$$

$$\begin{aligned} \Rightarrow \Delta V_{FB,pl} &= \int_0^{T_N} Q(x) \left(\frac{T_{Box}}{\epsilon_0 \epsilon_{ox}} + \frac{x}{\epsilon_0 \epsilon_N} \right) dx \\ &= Q \left(\frac{T_{Box}}{\epsilon_0 \epsilon_{ox}} + \frac{\hat{x}}{\epsilon_0 \epsilon_N} \right). \end{aligned} \quad (A-7)$$

Fig. 14. Calculated x - Q plots from Fig. 13.

APPENDIX B

THEORETICAL J VERSUS E_{Box} CALCULATION METHOD USING THE GATE-SENSING $V_{\text{FB,pl}}$ SHIFT

Case 1 (charges placed at the interface of nitride and top oxide)

$$J = \frac{dV_{\text{FB,pl}}}{dt} \frac{\epsilon_0 \epsilon_N \epsilon_{\text{ox}}}{T_{\text{Box}} \epsilon_N + T_N \epsilon_{\text{ox}}} \quad (\text{A-8})$$

$$E_{\text{Box}} = \frac{-\epsilon_N (\Delta V_{\text{FB,pl}} \cdot T_{\text{Tox}} \epsilon_N + V_G \cdot T_{\text{Box}} \epsilon_N + V_G \cdot T_N \epsilon_{\text{ox}})}{(T_{\text{Box}} \epsilon_N + T_N \epsilon_{\text{ox}})(T_{\text{Box}} \epsilon_N + T_{\text{Tox}} \epsilon_N + T_N \epsilon_{\text{ox}})}. \quad (\text{A-9})$$

Case 2 (charges placed at the interface of bottom oxide and nitride)

$$J = \frac{dV_{\text{FB,pl}}}{dt} \frac{\epsilon_0 \epsilon_{\text{ox}}}{T_{\text{Box}}} \quad (\text{A-10})$$

$$E_{\text{Box}} = \frac{-\Delta V_{\text{FB,pl}}}{T_{\text{Box}}} + \frac{(\Delta V_{\text{FB,pl}} - V_G) \epsilon_N}{T_{\text{Box}} \epsilon_N + T_{\text{Tox}} \epsilon_N + T_N \epsilon_{\text{ox}}}. \quad (\text{A-11})$$

Case 3 (charges placed at the center of nitride; (A-12) and (A-13), shown at the bottom of the page).

APPENDIX C

THEORETICAL $\Delta V_{\text{FB}} - t$ FOR DIFFERENT CHARGE VERTICAL LOCATIONS

The tunneling current through the bottom oxide O1 follows the FN equation:

$$J = \alpha E_1^2 \exp\left(\frac{-E_C}{E_1}\right). \quad (\text{A-14})$$

E_1 is the bottom oxide electric field, and

$$\alpha = \frac{q^3}{8\pi\hbar\phi_B} \frac{m}{m_s} \quad (\text{A-15})$$

$$E_C = 4\sqrt{2m_s} \frac{\phi_B^{3/2}}{3\hbar q} \quad (\text{A-16})$$

where q is the electron charge, \hbar is the reduced Planck constant, ϕ_B is the barrier height of oxide for electron tunneling, m is the mass of a free electron, and m_s is the effective mass of an electron in the oxide.

Channel-sensing Method

The bottom oxide electric field can be written as

$$E_1 = \frac{(-\Delta V_{\text{FB,ch}} + V_G) \epsilon_N}{T_{\text{Box}} \epsilon_N + T_{\text{Tox}} \epsilon_N + T_N \epsilon_{\text{ox}}}. \quad (\text{A-17})$$

$$J = \frac{dV_{\text{FB,pl}}}{dt} \frac{2\epsilon_0 \epsilon_N \epsilon_{\text{ox}}}{2T_{\text{Box}} \epsilon_N + T_N \epsilon_{\text{ox}}} \quad (\text{A-12})$$

$$E_{\text{Box}} = \frac{-\epsilon_N [\Delta V_{\text{FB,pl}} (2T_{\text{Tox}} \epsilon_N + T_N \epsilon_{\text{ox}}) + V_G (2T_{\text{Box}} \epsilon_N + T_N \epsilon_{\text{ox}})]}{(2T_{\text{Box}} \epsilon_N + T_N \epsilon_{\text{ox}})(T_{\text{Box}} \epsilon_N + T_{\text{Tox}} \epsilon_N + T_N \epsilon_{\text{ox}})} \quad (\text{A-13})$$

The $\Delta V_{\text{FB,ch}}$ can be derived by an iteration method

$$\Delta V_{\text{FB,ch}}^{(i+1)} = \Delta V_{\text{FB,ch}}^{(i)} + \delta t \times J \times \left(\frac{\varepsilon_0 \varepsilon_N \varepsilon_{\text{ox}}}{T_{\text{Tox}} \varepsilon_N + (T_N - x) \varepsilon_{\text{ox}}} \right)^{-1} \quad (\text{A-18})$$

where δt is the time interval used in the iteration, and x is the charge vertical location.

The $\Delta V_{\text{FB,ch}}$ for the charges placed at the interface of nitride and top oxide

$$\Delta V_{\text{FB,ch}}^{(i+1)} = \Delta V_{\text{FB,ch}}^{(i)} + \delta t \times J \times \left(\frac{\varepsilon_0 \varepsilon_{\text{ox}}}{T_{\text{Tox}}} \right)^{-1}. \quad (\text{A-19})$$

The $\Delta V_{\text{FB,ch}}$ for the charges placed at the interface of bottom oxide and nitride

$$\Delta V_{\text{FB,ch}}^{(i+1)} = \Delta V_{\text{FB,ch}}^{(i)} + \delta t \times J \times \left(\frac{\varepsilon_0 \varepsilon_N \varepsilon_{\text{ox}}}{T_{\text{Tox}} \varepsilon_N + T_N \varepsilon_{\text{ox}}} \right)^{-1}. \quad (\text{A-20})$$

The $\Delta V_{\text{FB,ch}}$ for the charges placed at the center of nitride

$$\Delta V_{\text{FB,ch}}^{(i+1)} = \Delta V_{\text{FB,ch}}^{(i)} + \delta t \times J \times \left(\frac{2\varepsilon_0 \varepsilon_N \varepsilon_{\text{ox}}}{2T_{\text{Tox}} \varepsilon_N + T_N \varepsilon_{\text{ox}}} \right)^{-1}. \quad (\text{A-21})$$

Gate-sensing Method

The E_1 in all cases are the same as the E_{Box} in Appendix B.

The $\Delta V_{\text{FB,pl}}$ for the charges placed at the interface of nitride and top oxide

$$\Delta V_{\text{FB,pl}}^{(i+1)} = \Delta V_{\text{FB,pl}}^{(i)} + \delta t \times J \times \left(\frac{\varepsilon_0 \varepsilon_N \varepsilon_{\text{ox}}}{T_{\text{Box}} \varepsilon_N + T_N \varepsilon_{\text{ox}}} \right)^{-1}. \quad (\text{A-22})$$

The $\Delta V_{\text{FB,pl}}$ for the charges placed at the interface of bottom oxide and nitride

$$\Delta V_{\text{FB,pl}}^{(i+1)} = \Delta V_{\text{FB,pl}}^{(i)} + \delta t \times J \times \left(\frac{\varepsilon_0 \varepsilon_{\text{ox}}}{T_{\text{Box}}} \right)^{-1}. \quad (\text{A-23})$$

The $\Delta V_{\text{FB,pl}}$ for the charges placed at the center of nitride

$$\Delta V_{\text{FB,pl}}^{(i+1)} = \Delta V_{\text{FB,pl}}^{(i)} + \delta t \times J \times \left(\frac{2\varepsilon_0 \varepsilon_N \varepsilon_{\text{ox}}}{2T_{\text{Box}} \varepsilon_N + T_N \varepsilon_{\text{ox}}} \right)^{-1}. \quad (\text{A-24})$$

Fig. 8 is calculated by assuming $m_s = 0.45$ m and $\phi_B = 3.1$ eV.

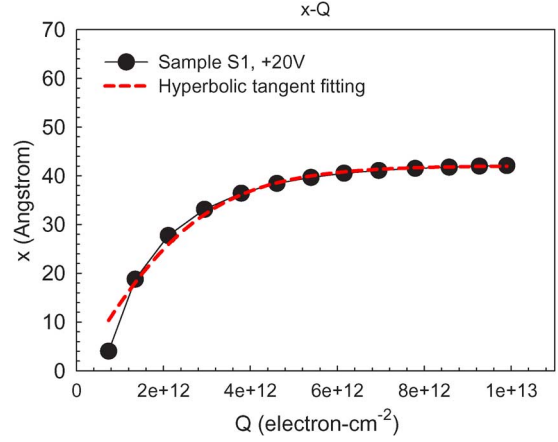


Fig. 15. Curve fitting of the experimental $x-Q$ plot (Fig. 3). Hyperbolic function is used to fit the experimental data. Next $x-Q$ curve is used to theoretically calculate $V_{\text{FB}} - t$ for both channel-sensing and gate-sensing.

APPENDIX D

THEORETICAL $\Delta V_{\text{FB}} - t$ CONSIDERING THE CHARGE INJECTION TRAJECTORY ($x-Q$ PLOT)

The tunneling current through the bottom oxide O1 follows the FN equation (A-14). At time zero, we assume that the ΔV_{FB} and charge vertical location (x) are zero, and they will increase as the time increases. During +FN programming, the Q and x always follow the relationship

$$x^{(i+1)} = A \tanh \left(\frac{Q^{(i)}}{L} \right). \quad (\text{A-25})$$

A and L are the fitting parameters.

This relationship is obtained from fitting the experimental $x-Q$ plot, as shown in Fig. 15.

Channel-sensing Method

The bottom oxide electric field is the same as (A-17). The $\Delta V_{\text{FB,ch}}$ can be derived by an iteration method

$$\Delta V_{\text{FB,ch}}^{(i+1)} = \Delta V_{\text{FB,ch}}^{(i)} + \delta t \times J \times \left(\frac{\varepsilon_0 \varepsilon_N \varepsilon_{\text{ox}}}{T_{\text{Tox}} \varepsilon_N + (T_N - x^{(i)}) \varepsilon_{\text{ox}}} \right)^{-1}. \quad (\text{A-26})$$

The total injected charges

$$Q^{(i+1)} = \frac{\Delta V_{\text{FB,ch}}^{(i+1)} \varepsilon_0 \varepsilon_N \varepsilon_{\text{ox}}}{T_{\text{Tox}} \varepsilon_N + (T_N - x^{(i)}) \varepsilon_{\text{ox}}}. \quad (\text{A-27})$$

Gate-sensing Method

The bottom oxide electric field is given by (A-28), shown at the bottom of the page.

$$E_1 = \frac{\varepsilon_N [\Delta V_{\text{FB,pl}}(x \varepsilon_{\text{ox}} - T_N \varepsilon_{\text{ox}} - T_{\text{Tox}} \varepsilon_N) + V_G (T_{\text{Box}} \varepsilon_N + x \varepsilon_{\text{ox}})]}{(T_{\text{Box}} \varepsilon_N + x \varepsilon_{\text{ox}})(T_{\text{Box}} \varepsilon_N + T_{\text{Tox}} \varepsilon_N + T_N \varepsilon_{\text{ox}})} \quad (\text{A-28})$$

The $\Delta V_{\text{FB,pl}}$ can be derived by an iteration method

$$\Delta V_{\text{FB,pl}}^{(i+1)} = \Delta V_{\text{FB,pl}}^{(i)} + \delta t \times J \times \left(\frac{\varepsilon_0 \varepsilon_N \varepsilon_{\text{ox}}}{T_{\text{Box}} \varepsilon_N + x^{(i)} \varepsilon_{\text{ox}}} \right)^{-1}. \quad (\text{A-29})$$

The total injected charges

$$Q^{(i+1)} = \frac{\Delta V_{\text{FB,pl}}^{(i+1)} \varepsilon_0 \varepsilon_N \varepsilon_{\text{ox}}}{T_{\text{Box}} \varepsilon_N + x^{(i)} \varepsilon_{\text{ox}}}. \quad (\text{A-30})$$

Fig. 10 is calculated by assuming that $m_s = 0.485$ m and that $\phi_B = 3.1$ eV.

REFERENCES

- [1] K. Kim, "Technology for sub-50 nm DRAM and NAND flash manufacturing," in *IEDM Tech. Dig.*, 2005, pp. 333–336.
- [2] E. Suzuki, Y. Hayashi, K. Ishii, and T. Tsuchiya, "Traps created at the interface between the nitride and the oxide on the nitride by thermal oxidation," *Appl. Phys. Lett.*, vol. 42, no. 7, pp. 608–610, Apr. 1983.
- [3] T. Ishida, Y. Okuyama, and R. Yamada, "Characterization of charge traps in metal–oxide–nitride–oxide–semiconductor (MONOS) structures for embedded Flash memories," in *Proc. IEEE 44th Annu. Int. Reliab. Phys. Symp.*, 2006, pp. 516–522.
- [4] H. T. Lue, Y. H. Shih, K. Y. Hsieh, R. Liu, and C. Y. Lu, "A transient analysis method to characterize the trap vertical location in nitride-trapping devices," *IEEE Electron Device Lett.*, vol. 25, no. 12, pp. 816–818, Dec. 2004.
- [5] A. Arreghini, F. Driussi, D. Esseni, L. Selmi, M. J. van Duuren, and R. van Schaijk, "Experimental extraction of the charge centroid and of the charge type in the P/E operation of SONOS memory cells," in *IEDM Tech. Dig.*, 2006, pp. 1–4.
- [6] H. T. Lue, P. Y. Du, S. Y. Wang, E. K. Lai, K. Y. Hsieh, R. Liu, and C. Y. Lu, "A novel gate-sensing and channel-sensing transient analysis method for real-time monitoring of charge vertical location in SONOS-type devices and its applications in reliability studies," in *Proc. IEEE 45th Annu. Int. Reliab. Phys. Symp.*, 2007, pp. 177–184.
- [7] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. New York: Wiley, 1983.
- [8] H. T. Lue, S. Y. Wang, E. K. Lai, Y. H. Shih, S. C. Lai, L. W. Yang, K. C. Chen, J. Ku, K. Y. Hsieh, R. Liu, and C. Y. Lu, "BE-SONOS: A bandgap engineered SONOS with excellent performance and reliability," in *IEDM Tech. Dig.*, 2005, pp. 547–550.



Hang-Ting Lue (M'05) was born in Hsinchu, Taiwan, R.O.C., in 1975. He received the B.S. and M.S. degrees in physics from National Tsing-Hua University, Hsinchu, in 1997 and 1999, respectively, and the Ph.D. degree in electrical engineering from National Chiao-Tung University, Hsinchu, in 2002.

He has been with the Emerging Central Laboratory, Macronix International Company, Ltd., Hsinchu, since 2003, where his current research is engaged in developing novel Flash memory devices and related theoretical modeling. From 2004, he has

published more than ten papers in the premier semiconductor conferences, including IEDM, VLSI, and IRPS, and more than 30 technical papers in IEEE journal/letters.

Dr. Lue is currently a committee member of Very Large Scale Integration Technology, Systems, and Applications.



Szu-Yu Wang was born in Kaohsiung, Taiwan, R.O.C., in 1974. He received the B.S. and M.S. degrees in material science and engineering from National Tsing-Hua University, Hsinchu, Taiwan, in 1996 and 1998, respectively, where he is currently working toward the Ph.D. degree in electronic engineering.

Since 2000, he has been with the Technology Development Center, Macronix International Company, Ltd., Hsinchu, where his current work is focused on dielectric characterizations in nitride-trapping Flash memories.



Erh-Kun Lai received the M.S. degree in electrical engineering from National Sun Yet-sen University, Kaohsiung, Taiwan, R.O.C., in 1997.

He has been with the Emerging Central Laboratory, Macronix International Company, Ltd., Hsinchu, Taiwan, since 1997. His current studies include 3-D charge-trapping nonvolatile-memory devices, resistive RAM, programmable conductor RAM, and process integration.



Tiao-Yuan Huang (S'78–M'78–SM'88–F'95) received the Ph.D. degree from the University of New Mexico, Albuquerque, in 1981.

After working in the U.S. semiconductor industry for 14 years, he returned to his native country, Taiwan, and has since been a Professor with the National Chiao-Tung University, Hsinchu, Taiwan, R.O.C.

Dr. Huang was elected as an IEEE fellow in 1995 for his invention and demonstration of fully overlapped lightly doped-drain MOS transistors.



Kuang-Yeu Hsieh was born in Tainan, Taiwan, R.O.C., in 1958. He received the B.S. degree in physics from National Tsing-Hua University, Hsinchu, Taiwan, in 1980, the M.S. degree in materials science from National Sun Yet-sen University, Kaohsiung, Taiwan, in 1985, and the Ph.D. degree in materials science from North Carolina State University, Raleigh, in 1989.

Before he joined the Emerging Central Laboratory, Macronix International Company, Ltd., Hsinchu, where he is currently with since 2001, he

was an Associate Professor with the Institute of Materials Science, National Sun Yet-sen University for nine years. Currently, his research involves developing new nonvolatile-memory devices and exploring new material for the next-generation nonvolatile memory. His research interests include molecular-beam-epitaxy thin-film growth, characterization of material, solid-state physics, IC fabrication, and optoelectronic materials.



Pei-Ying Du was born in Taipei, Taiwan, R.O.C., in 1982. She received the B.S. degree in engineering and system science from National Tsing-Hua University, Hsinchu, Taiwan, in 2004. She is currently working toward the Ph.D. degree at the Institute of Electronics, National Chiao-Tung University, Hsinchu.

She has been with the Emerging Central Laboratory, Macronix International Company, Ltd., Hsinchu, since 2006, where her current research is engaged in reliability studies of nitride-trapping Flash memories.



Rich Liu (SM'00) received the B.S. degree in nuclear engineering from Hsing Hua University, Hsinchu, Taiwan, R.O.C., the M.S. degree in physics from Ohio State University, Columbus, and the Ph.D. degree in materials science from the University of Illinois, Urbana-Champaign.

In 1979, he was with Bell Laboratories, Murray Hill, NJ, as a member of Technical Staff. Since then, he has engaged in research and advanced development of e-beam lithography systems, electronic materials and CMOS devices, interconnect and integration, and optical MEMS, and over the years, he became a Distinguished Member of Technical Staff, a Supervisor, a Technical Manager, and a Senior Manager with Bell Laboratories, Lucent Technologies. He has been with the Macronix International Company, Ltd., since 2001, and is currently the Chief Scientist for the Macronix Innovation R&D Center and the Executive Director of the Emerging Central Laboratory.

Dr. Liu served as an Editor of the *IEEE ELECTRON DEVICE LETTERS* and served in numerous IEEE-sponsored conference committees.



Chih-Yuan Lu (M'78-SM'84-F'95) received the B.S. degree from National Taiwan University, Taipei, Taiwan, R.O.C., in 1972, and the Ph.D. degree in physics from Columbia University, New York, NY, in 1977.

He was a Professor with National Chiao-Tung University and was with AT&T Bell Laboratories from 1984 to 1989. He later joined the Electronics Research and Service Organization/Industrial Technology Research Institute (ITRI), Hsinchu, Taiwan, as a Deputy General Director, responsible for the MOEA Grand Submicron Project, in 1989. This project successfully developed Taiwan's first 8-in manufacturing technology with high-density DRAM/SRAM. In 1994, he became the Cofounder of the Vanguard International Semiconductor Corporation, which is a spin-off memory IC Company from ITRI's Submicron Project. He was the Vice President of Operation, the Vice-President of R&D, and later the President from 1994 to 1999. He is currently the Chairman and Chief Executive Officer of Ardentec Corporation, a VLSI testing service company. He also serves the Emerging Central Laboratory, Macronix International Company, Ltd. (MXIC), Hsinchu, as a Senior Vice-President/CTO. He led the MXIC's technology development team to successfully achieve the state-of-the-art nonvolatile-memory technology and is now responsible for the MXIC's overall memory operation. He has published more than 100 papers and has been granted 123 international patents.

Dr. Lu was elected as a Fellow of the American Physical Society. He is the recipient of the highest honor—the National Science and Technology Achievement Award by the Prime Minister of China—due to his leadership and achievement in this Submicron Project. He is also the recipient of the IEEE Millennium Medal, and the most prestigious semiconductor R&D Award in Taiwan, from Pan Wen Yuan Foundation.