Numerical Simulation of Bottom Oxide Thickness Effect on Charge Retention in SONOS Flash Memory Cells

Shaw-Hung Gu, Chih-Wei Hsu, Tahui Wang, *Senior Member, IEEE*, Wen-Pin Lu, Yen-Hui Joseph Ku, and Chih-Yuan Lu

***Abstract*—In this paper, bottom-oxide thickness (T_{bo}) and program/erase stress effects on charge retention in SONOS Flash memory cells with FN programming are investigated. Utilizing a numerical analysis based on a multiple electron-trapping model to solve the Shockley–Read–Hall rate equations in nitride, we simulate the electron-retention behavior in a SONOS cell with T_{bo} from 1.8 to 5.0 nm. In our model, the nitride traps have a continuous energy distribution. A series of Frenkel–Poole (FP) excitation of trapped electrons to the conduction band and electron recapture into nitride traps feature the transitions between the conduction band and trap states. Conduction band electron tunneling via oxide traps created by high-voltage stress and trapped electron direct tunneling through the bottom oxide are included to describe various charge leakage paths. We measure the nitride-charge leakage current directly in a large-area device for comparison. This paper reveals that the charge-retention loss in a high-voltage stressed cell, with a thicker bottom oxide (5 nm), exhibits two stages. The charge-leakage current is limited by oxide trap-assisted tunneling in the first stage and, then, follows a $1/t$ time dependence due to the FP emission in the second stage. The transition time from the first stage to the second stage is related to oxide trap-assisted tunneling time but is prolonged by a factor.**

***Index Terms*—Oxide thickness, positive oxide charge-assisted tunneling, Shockley–Read–Hall (SRH) rate equation, SONOS retention mechanisms.**

I. Introduction

RECENTLY, considerable research efforts have been made to study nitride-trap storage Flash memories for their lower operation voltage, simpler fabrication process, and better scaling capability as compared to conventional floating gate Flash memory [1]–[4]. The nitride storage cells have evolved into two types: The first one has uniform charge storage in a nitride layer, such as conventional SONOS, and the second one utilizes localized charge storage at the source/drain junctions, such as NROM [1] or Nbit technology [4]. Since the conventional SONOS cell employs FN tunneling for program/erase (P/E), a thinner bottom oxide is needed as compared to NROM. Its data retention, thus, imposes an intrinsic reliability constraint due to significant direct-tunneling current through an ultrathin bottom oxide [5]–[8].

Several charge-retention-loss mechanisms in SONOS devices have been reported. Lundkvist *et al.* showed that trapped-charge direct tunneling leads to charge loss at room temperature [8], and later, they accounted for the increased decay rate of charge loss at elevated temperatures with a thermal-enhanced charge-emission model [9], [10]. Lehovec and Fedotowsky delivered a simple analytical retention model through the Frenkel–Poole (FP) release of electrons from monoenergetic nitride traps [11]. Williams and Beguwala regarded charge loss as direct tunneling of the charges out of the nitride together with nitride-charge migration by a series of emission-capture events [12]. Recently, White and coworkers have combined trap-to-band tunneling and thermal excitation to reproduce measured device retention data [13], [14]. Although a variety of models have been proposed, they have some deficiencies in common. For example, all the models are applicable only to ultrathin bottom oxides (1.5 $\sim$ 2.5 nm). However, to improve data retention and to minimize gate disturb, a thicker bottom oxide is usually employed in today's SONOS cells [15], [16]. With a larger oxide thickness, oxide traps created by P/E stress becomes important and should be taken into account in a charge-loss model.

In this paper, we develop a numerical approach to solve a set of rate equations governing the electron FP emission and retrapping in the nitride and charge leakage through the bottom oxide [17]–[21]. Comprehensive charge-loss mechanisms based on direct tunneling for trapped electrons and positive charge (oxide hole trap)-assisted tunneling (PCAT) [17] for conduction-band electrons in nitride are formulated. The effects of bottom-oxide thickness (T_{bo}) and stress-created oxide traps on charge retention are measured and simulated. The dominant leakage mechanisms for different T_{bo} are identified.

II. Retention-Loss Simulation Model

The samples used in this paper are n-channel SONOS cells consisting of a top oxide of $T_{to} = 9$ nm and a silicon nitride of $T_n = 6$ nm. The T_{bo} ranges from 1.8 to 5 nm. The measurement data for the $T_{bo} = 1.8$ nm device are quoted from [5]

Manuscript received March 13, 2006; revised August 16, 2006. The work of S.-H. Gu and T. Wang was supported by the National Science Council (NSC), Taiwan, R.O.C. under Contract NSC 92-2215-E-009-004. The review of this paper was arranged by Editor S. Kimura.

S.-H. Gu, C.-W. Hsu, and T. Wang are with the Department of Electronics Engineering, National Chiao-Tung University, Hsinchu 300, Taiwan, R.O.C. (e-mail: twang@cc.nctu.edu.tw).

W.-P. Lu, Y.-H. J. Ku, and C.-Y. Lu are with the Macronix International Company Ltd., 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/TED.2006.887219

0018-9383/$25.00 © 2007 IEEE

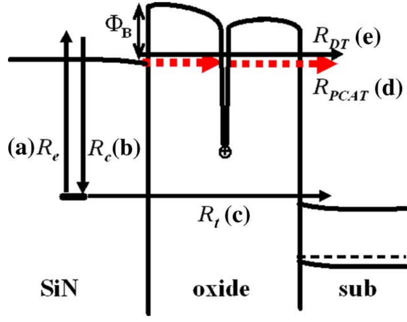


Fig. 1. Illustration of nitride-electron leakage paths and electron transitions between the conduction band and trap states. (a) FP excitation of trapped electrons to the Si/N conduction band. (b) Free electrons recaptured by nitride traps. (c) Direct tunneling of nitride-trapped electrons to Si substrate. (d) Free electrons via PCAT to Si substrate. (e) Direct tunneling of free electrons to Si substrate.

and other oxide thickness results are measured in this paper. FN tunneling is employed for programming and erase. This uniform FN injection excludes the possibility of stored charge lateral migration in the nitride.

A. Nitride-Charge Dynamics and Loss Mechanisms

Charge transitions between the conduction band and trap states in the nitride and stored-charge leakage paths are illustrated in the energy-band diagram in Fig. 1. The change of conduction band and trapped electron densities in the nitride is described by the following two-coupled rate equations [18]–[20]:

$$\frac{dn_t(x,\phi,t)}{dt} = R_c(x,\phi,t)\cdot n_c(t) - R_e(\phi)\cdot n_t(x,\phi,t) - R_t(x,\phi)\cdot n_t(x,\phi,t) \quad (1)$$

$$\frac{dn_c(t)}{dt} = \frac{1}{T_n}\iint \left[R_e(\phi)\cdot n_t(x,\phi,t) - R_c(x,\phi,t)n_c(t)\right] dx d\phi - (R_{\mathrm{PCAT}} + R_{\mathrm{DT}})\cdot n_c(t) \quad (2)$$

where $n_t(x,\phi,t)$ (per cubic centimeter per electronvolt) is the density of occupied nitride traps as a function of a distance to the SiN/bottom-oxide interface (x), trap energy ϕ, and retention time t, respectively. The density of electrons in the nitride conduction band is denoted by n_c (per cubic centimeters), which is assumed to be uniform in x, because the nitride layer is sufficiently thin. R_e, R_c, R_t, R_{PCAT}, and R_{DT} are the rate coefficients for FP electron emission from nitride traps (Fig. 1, path a), free-electron capture into nitride traps (Fig. 1, path b), trapped electron direct tunneling to the substrate (Fig. 1, path c), conduction-band-electron escape through PCAT (Fig. 1, path d) and via direct tunneling (Fig. 1, path e), respectively. The FP excitation of electrons from nitride traps to the conduction band is [11]

$$R_e(\phi) = \nu_e \cdot \exp\left(\frac{\beta\sqrt{E_n} - \phi}{kT}\right) \quad (\text{unit} : \text{s}^{-1}) \quad (3)$$

where

$$\nu_e = \mathrm{N_{cn}}\cdot \upsilon_{\mathrm{th}}\cdot \sigma_n. \quad (\text{unit} : \text{s}^{-1}) \quad (3a)$$

$$\tau_e = R_e^{-1} \quad (\text{unit} : \text{s}). \quad (3b)$$

The prefactor ν_e is often referred to as the "attempt-to-escape" frequency for emission and can be expressed as (3a), where N_{cn} is the effective density of states in the nitride conduction band, υ_{th} is the thermal velocity, and σ_n is the nitride-trap capture cross section. The FP constant is β, and E_n is the average electric field in the nitride. Moreover, $\tau_e(\phi)$ is the electron-emission time from the trap energy of ϕ. The free-electron capture-rate coefficient is

$$R_c(x,\phi,t) = \upsilon_{\mathrm{th}}\cdot\sigma_n\cdot(N_t - n_t) \quad (\text{unit} : \text{s}^{-1}\text{eV}^{-1}) \quad (4)$$

where N_t is the nitride-trap density per-unit-trap energy and $(N_t - n_t)$ is the amount of available traps for free-electron recapture. The direct-tunneling rate for nitride-trapped electron is

$$R_t(x,\phi) = \nu_t\cdot\exp(-\alpha_{\mathrm{ox}}T_{\mathrm{bo}}) \cdot\exp(-\alpha_n x) \quad (\text{unit} : \text{s}^{-1}) \quad (5)$$

$$\alpha_{\mathrm{ox}} = \frac{2\sqrt{2m^*_{\mathrm{ox}}q(\Phi_B+\phi)}}{\hbar}; \quad \alpha_n = \frac{2\sqrt{2m^*_n q\phi}}{\hbar} \quad (5a)$$

where m^*_{ox} and m^*_n are electron tunneling mass in the oxide and in the nitride, respectively. Other variables have their usual definitions.

As for conduction band electron leakage paths, positive oxide charge-assisted tunneling [17] and direct tunneling through the bottom oxide are considered. Charge leakage via top oxide is not considered for a relatively large top oxide thickness. The built-in electric field in the ONO stack is neglected for simplification. Electron tunneling from nitride traps to surface traps in Si band gap is also neglected. To simplify the evaluation of PCAT, a concept of the most favorable trapped-charge position is employed [22]. With an oxide charge site at this favorable position, the tunneling probability from the conduction band to the trap site is the same as that from the trap site to the substrate conduction band. Such feature of the most efficient tunneling trap location leads to the following equation:

$$R_{\mathrm{PCAT}} = N_{\mathrm{ox}}\cdot\upsilon_{\mathrm{th}}\cdot\sigma_{\mathrm{ox}}\cdot\mathrm{P_{PCAT}} \quad (\text{unit} : \text{s}^{-1}) \quad (6)$$

$$\tau_{\mathrm{PCAT}} = R^{-1}_{\mathrm{PCAT}} \quad (\text{unit} : \text{s}) \quad (6a)$$

where N_{ox} is the stress-induced oxide hole-trap (positively charged) density, P_{PCAT} is the tunneling probability from the nitride conduction band to the substrate, and τ_{PCAT} is the electron tunneling time via PCAT. The way to calculate the tunneling probability P_{PCAT} is the same as in [17]. A one-dimensional Coulombic potential caused by a positive trapped charge is included in the electron-tunneling barrier for the Wentzel–Kramers–Brillouin (WKB) approximation, i.e.,

$$\Phi_{\mathrm{coul}}(x) = \frac{q}{4\pi\varepsilon_{\mathrm{ox}}|x - x_h|} \quad (7)$$

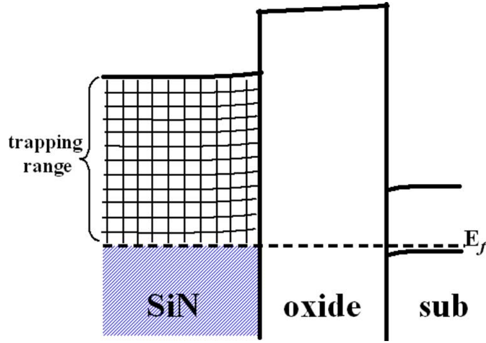


Fig. 2. Illustration of the discretization scheme in trap energy and space for numerical simulation.

where $\varepsilon_{\rm ox}$ is permittivity of oxide and x_h denotes the location of a trapped positive charge from SiN/bottom-oxide interface. Besides, the direct tunneling rate can be deduced from the supply function

$$R_{\rm DT} = \frac{\upsilon_{\rm th}}{T_n} \cdot \exp(-\alpha_{\rm ox} T_{\rm bo}) \qquad (\text{unit} : \text{s}^{-1}) \tag{8}$$

with $\alpha_{\rm ox}$ (5a) evaluated at the conduction-band edge. With the above rate coefficients determined, the nitride-charge leakage current J can be written as

$$J = q \iint R_t(x,\phi,t) \cdot n_t(x,\phi,t) dx d\phi + q \cdot R_{\rm PCAT} \cdot n_c(t) \cdot T_n + q \cdot R_{\rm DT} \cdot n_c(t) \cdot T_n. \tag{9}$$

The first term on the right-hand side is the ensemble of out-tunneling electrons from nitride traps. The second and the third terms represent charge loss from the conduction band via PCAT and direct tunneling (DT). It should be mentioned that in a P/E stressed device with a thicker oxide, direct tunneling is much smaller than PCAT and (9) reduces to

$$J = q \cdot R_{\rm PCAT} \cdot n_c(t) \cdot T_n. \tag{10}$$

Several assumptions concerning the nitride traps in our model should be mentioned. First, we assume that the nitride traps have a continuous distribution in the bandgap and the trap density decreases exponentially with trap energy, i.e., $N_t(\phi) = N_{\rm to} \cdot \exp(-\phi/\lambda)$, where $N_{\rm to}$ is set to be 3.0×10^{19} cm$^{-3} \cdot$ eV^{-1} and λ is the tail parameter (0.85 eV). This exponential trap-energy distribution is often used for amorphous materials [18], [23], [24]. Second, the nitride traps below the Fermi-level are occupied in a fresh device in thermal equilibrium and, thus, are not available for the storage of injected charges. Third, nitride traps are considered to be isolated from each other and direct electron transitions among nitride traps are not taken into account.

B. Numerical Simulation Method

The two nonlinear rate [(1) and (2)] are numerically solved by using an implicit backward Euler method coupled with Newton iteration [25]. In our numerical calculation, the nitride traps are discretized into equidistant segments in energy and in space, as shown in Fig. 2, transforming the differential equations into numerous sets of mutually related equations. The transient response was computed for each segment separately; thus, providing an insight into physical events of trap states in charge loss process. The parameters used in our model are given in Table I. To simulate the retention behavior, the total amount of stored electrons in the nitride is determined by measured threshold voltage (V_t) shift after programming. The initial electron distribution in the nitride layer is obtained as the injected charges attain the dynamic balancing between the nitride conduction band and the trap states by setting all leakage paths equal to zero. Once the initial condition is determined, the conduction-band-electron and trapped electron densities at the subsequent instant can be evaluated step by step.

TABLE I
PARAMETERS USED FOR THE NUMERICAL SIMULATION

Parameters	Value
$\varepsilon_{\rm ox}$	$3.9\varepsilon_o$
$\varepsilon_{\rm N}$	$7\varepsilon_o$
$m_{\rm ox}$	$0.42m_o$ [27]
m_n	$0.42m_o$ [27]
Φ_B	1.05 eV [13,14]
N_{t0}	3.0×10^{19} cm^{-3}·eV^{-1}
σ_n	5×10^{-13} cm^2 [13,14]
$\sigma_{\rm ox}$	1×10^{-14} cm^2
$\upsilon_{\rm th}$	10^7 cm/s
ν_e	5×10^{15} s^{-1}
ν_t	1.4×10^{14} s^{-1}

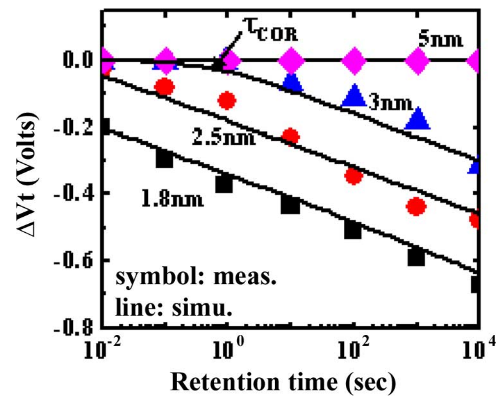


Fig. 3. Program state V_t retention behavior for different $T_{\rm bo}$ at $T = 25$ °C. All devices are programmed to an identical threshold-voltage window of 1 V. The symbols represent measured data and the lines are simulated results. The corner time ($\tau_{\rm COR}$) is indicated.

III. RESULTS AND DISCUSSION

A. Bottom-Oxide Blocking Effect

To investigate the $T_{\rm bo}$ effect, the precycling-retention characteristics of four SONOS devices with different $T_{\rm bo}$ (1.8, 2.5, 3, and 5 nm) are compared in Fig. 3. All the devices are programmed to have an identical threshold-voltage window of 1 V. Good agreement between the measured (solid lines) and simulated data (symbols) confirms the validity of our simulation model. The program state V_t (for example, the 3-nm oxide cell) remains almost unchanged for a certain

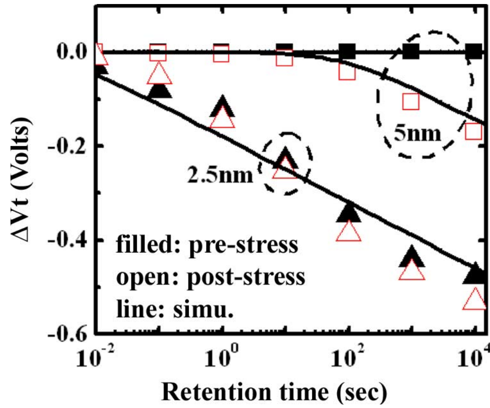


Fig. 4. Prestress and poststress V_t loss for $T_{\rm bo} = 2.5$ and 5 nm. $T = 25$ °C. FN stress was performed at Eox $= -11$ MV/cm for 100 s.

period of time and, then, decreases with a $\log(t)$ dependence. The charge loss is retarded with increasing $T_{\rm bo}$. This feature can be well explained by a tunneling front model [26]. The corner time $(\tau_{\rm COR})$ in Fig. 3 corresponds to trapped electron direct-tunneling time [Fig. 1, path (c)]. The role of high-voltage stress created oxide traps in charge retention is examined in Fig. 4 for $T_{\rm bo} = 5$ and 2.5 nm. The two devices are subjected to a negative FN stress. As opposed to the thinner oxide (2.5 nm) device, the 5-nm oxide cell exhibits an apparent stress effect. The poststress V_t loss in the 5-nm oxide device exhibits an initial delay and, then, declines with a $\log(t)$ time dependence. Our calculation in the next section will show that the dominant retention-loss mechanism in the poststress 5-nm oxide cell is the FP emission followed by PCAT.

B. Poststress Two-Stage Retention Loss

To explain the above poststress retention-loss behavior in the 5-nm oxide device, we measure the stress-induced leakage current [i.e., (9)] in a large-area device (500×500 μm) directly at $T = 25$ °C. The programming window is 3 V, and the measurement gate voltage is 0 V. The stress-induced leakage current (gate current) and the evolution of the retention V_t are shown in Fig. 5(a) and (b). One of the samples in Fig. 5(a) is stressed lightly, and the other is stressed more heavily. The symbols in Fig. 5 denote measured data, and the solid lines represent simulation result. In the simulation, we use the positive trapped-charge (hole trap) density as a fitting parameter. The hole-trap density is 4.5×10^{17} cm^{-3} for the lightly stressed sample and 3×10^{18} cm^{-3} for the heavily stressed sample. The temperature dependence of V_t retention loss is shown in Fig. 6 for $T = 25$ °C and 85 °C. The higher bake temperature yields a larger V_t loss, which is in agreement with the FP model. Trap anneal effect should be considered as bake temperature further increases. It should be noted that the nitride-charge leakage current exhibits two stages (Fig. 5). In the first stage $(t < \tau_{\rm COR})$, a dc-like characteristic, termed as current-blocking effect, is observed. The current-blocking effect persists for a longer time in the lightly stressed sample. After $\tau_{\rm COR}$, a $1/t$ transient decay in the leakage current is observed, which accounts for the $\log(t)$ dependence of the V_t loss in Fig. 5(b). The evolution of nitride conduction-band-electron density (n_c) with retention

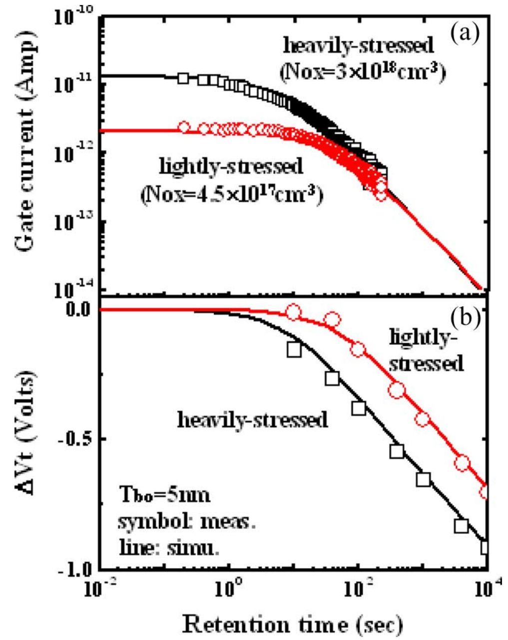


Fig. 5. (a) Measured stress-induced leakage current (gate current) in a large-area device (500×500 μm) at $V_g = 0$ V and $T = 25$ °C. The lightly and heavily stressed conditions are at Eox $= -11$ MV/cm for 1 s and 1000 s, respectively. Both devices are programmed to an identical threshold-voltage window of 3 V. (b) Corresponding V_t retention loss for the two stressed cells.

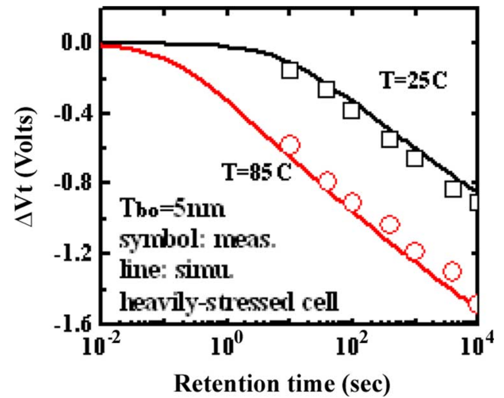


Fig. 6. Measured and simulated V_t loss versus retention time in the heavily stressed sample at $T = 25$ °C and 85 °C.

time is plotted in Fig. 7. The prestress one is also shown as a reference. In the first stage, n_c remains almost the same as the prestress value, because the oxide leakage current is limited by the amount of stress-created oxide traps. The conduction-band-electron density begins to have an apparent decrease in the second stage. The simulated nitride-trap electron occupation factor (f_t) versus trap energy at different retention times are shown in Fig. 8. The FP emission front $(\phi_{\rm fro})$, defined as the trap energy with $f_t = 0.5$, is displayed in Fig. 9 as a function of retention time. In the first stage, since the nitride-charge leakage current is limited by oxide trap-assisted tunneling, the FP emission front stays around 0.8 eV. In the second stage, the oxide trap-assisted tunneling is no longer a limiting mechanism, and the leakage current is dictated by the FP emission. As a consequence, the FP emission front moves downward in the nitride bandgap with a constant speed in a $\log(t)$ scale.

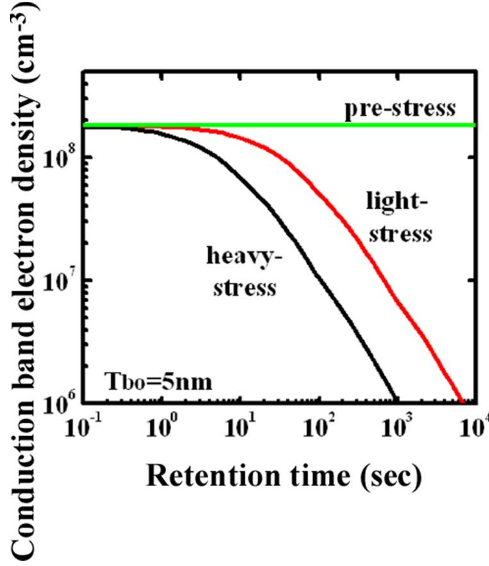


Fig. 7. Calculated conduction-band-electron density versus retention time before stress and after light stress and heavy stress, respectively. $T = 25\ ^\circ\mathrm{C}$.

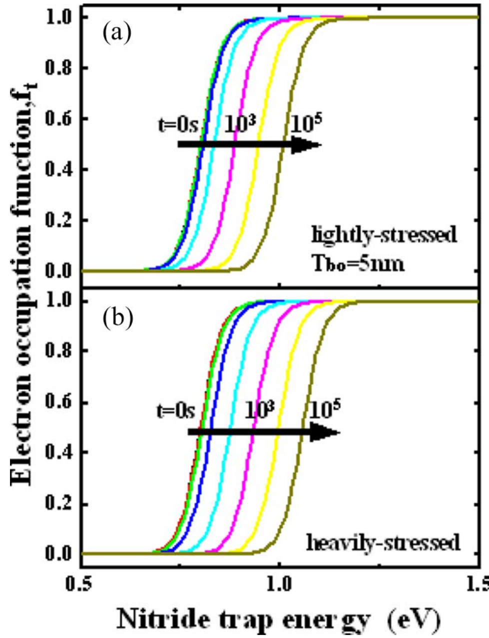


Fig. 8. Simulated electron occupation factor of nitride traps versus trap energy at different retention times. (a) Lightly stressed device. (b) Heavily stressed device.

For the heavily stressed sample, the oxide leakage current is higher in the first stage and the corner time τ_{COR} is shorter. To further elaborate on the τ_{COR}, we use a simplified picture by assuming that the nitride-charge transitions mainly occur between conduction band and the states in the vicinity of the FP emission front (in an energy range of ∼kT), as illustrated in Fig. 10. $\tau_e(\phi_{\mathrm{fro}})$ and $\tau_c(\phi_{\mathrm{fro}})$ represent electron emission time and capture time between the conduction band and the emission front. It should be pointed out that in our measurement period, τ_{PCAT} is much longer than $\tau_c(\phi_{\mathrm{fro}})$ that the conduction-band-electron density is mainly determined by τ_e, τ_c, and the trapped-charge density at the emission front $n_t(\phi_{\mathrm{fro}})$. Therefore, the

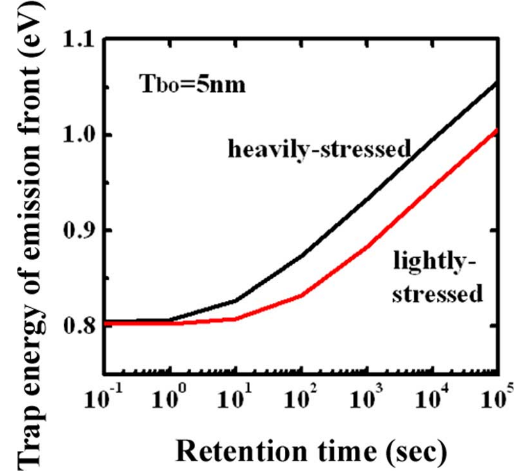


Fig. 9. Temporal evolution of the FP emission front, which is defined as the trap energy for the occupation factor in Fig. 8 equal to 0.5.

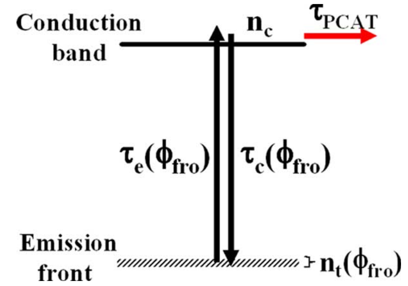


Fig. 10. Illustration of electron emission and capture between the conduction band and trap states around the FP emission front in an energy range of kT.

nitride-charge leakage current in the second stage can be readily obtained in the following [24]:

$$J \propto \frac{n_c}{\tau_{\mathrm{PCAT}}} \propto \frac{n_t(\phi_{\mathrm{fro}})}{\tau_{\mathrm{PCAT}}} \cdot \frac{\tau_c(\phi_{\mathrm{fro}})}{\tau_e(\phi_{\mathrm{fro}}) + \tau_c(\phi_{\mathrm{fro}})} = \frac{n_t(\phi_{\mathrm{fro}})}{\tau_{\mathrm{eff}}} \tag{11}$$

and

$$\tau_{\mathrm{eff}} = \frac{\tau_e(\phi_{\mathrm{fro}}) + \tau_c(\phi_{\mathrm{fro}})}{\tau_c(\phi_{\mathrm{fro}})} \cdot \tau_{\mathrm{PCAT}} \tag{12}$$

where τ_{eff} is the effective time for nitride-trapped charge at the emission front to escape from the ONO film. Fig. 11 shows our calculated τ_{eff} versus retention time. Three points are worth noting: 1) Since $n_t(\phi)$ has a pretty uniform distribution in our measurement period, J [Fig. 5(a)] and τ_{eff} (Fig. 11) have an inverse time-dependence from (11). 2) The corner time τ_{COR} in Figs. 4 and 5 is equal to τ_{PCAT} multiplied by a factor $(\tau_e + \tau_c)/\tau_c$, which is determined by the initial FP emission front. Fig. 12 shows the calculated τ_{COR} and τ_{PCAT} versus Nox. For a larger program-state V_t, the initial ϕ_{fro} is smaller and so is the factor $(\tau_e + \tau_c)/\tau_c$. Thus, the τ_{COR} becomes shorter as the program V_t window increases. This feature is consistent with our measurement result. It also should be mentioned that the neglect of conduction-band-electron recapture (i.e., $\tau_c = \infty$) would lead to $(\tau_e + \tau_c)/\tau_c = 1$ and an erroneous interpretation of $\tau_{\mathrm{COR}} = \tau_{\mathrm{PCAT}}$. 3) At a sufficiently long retention time, Fig. 11 shows that τ_{eff} is actually equal to retention time t. Thus,

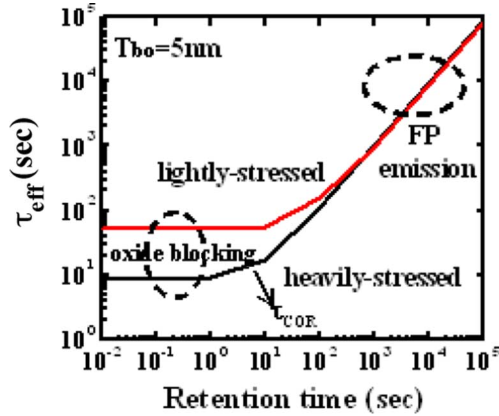


Fig. 11. Calculated effective time for nitride-trapped charge at the FP emission front to escape from the ONO film.

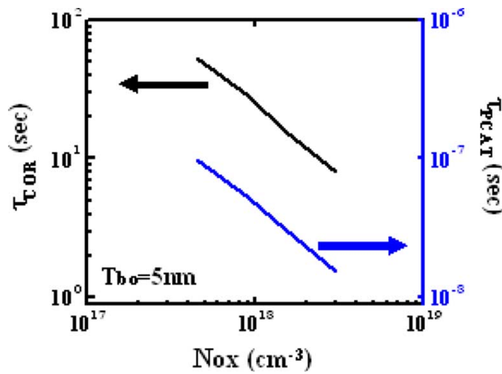


Fig. 12. Corner time (τ_{COR}) and positive oxide charge-assisted electron tunneling time (τ_{PCAT}) versus positively charged oxide trap density (Nox).

(11) becomes $J \propto n_t/t$ without regard to stress condition. This result is confirmed by our measurement in Fig. 5(a) that both samples (heavily stressed and lightly stressed) have the same leakage current in the second stage. A similar expression for the nitride leakage current in the second stage was published in [4].

C. Leakage Component Separation

As aforementioned, the role of PCAT in charge loss has been substantiated for $T_{\mathrm{bo}} = 5$ nm. On the other side, charge direct tunneling from nitride-trap states has been shown to be the dominant loss mechanism for ultrathin oxides at room temperature [8]. It is interesting to investigate the transition between these two mechanisms with respect to oxide thickness. Since they are difficult to be separated experimentally, a numerical approach is performed. Positive oxide charges are placed at the most efficient PCAT position for different oxide thickness. Fig. 13 shows the percentage of the V_t loss from the two components versus oxide thickness at $T = 25$ °C. For thinner oxides ($<$ 3 nm), nitride-charge loss is mainly via trapped-charge direct tunneling [Fig. 13, process (c)]. The PCAT becomes dominant for oxide thickness exceeding 3.7 nm.

IV. CONCLUSION

A numerical method to simulate the retention characteristics in SONOS Flash memories has been developed. The dominant

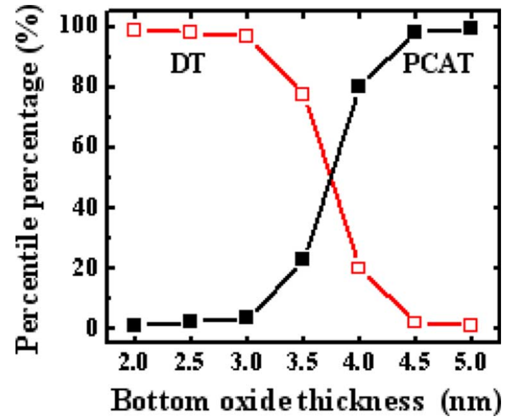


Fig. 13. Percentage of PCAT [process (d)] and DT [process (c)] caused retention V_t loss as a function of T_{bo} at $T = 25$ °C. Positively charged oxide traps with a density of 3×10^{18} cm^{-3} are placed at the most favorable position of PCAT. The dominant charge-loss mechanism changes from DT to PCAT as oxide thickness increases.

V_t loss mechanism is identified for different oxide thickness. The V_t loss in a SONOS Flash exhibits two stages. For a thinner bottom oxide, the transition time between the two stages is equal to the trapped-charge direct-tunneling time. For a thicker oxide, the transition time is related to hole-trap-assisted tunneling time but is prolonged by a factor. No matter of thinner or thicker oxides, the V_t retention loss in the second stage obeys a $\log(t)$ dependence. The dominant mechanism is trapped-charge direct tunneling for thinner oxides and the FP emission followed by positive oxide charge-assisted tunneling can well explain the observed retention characteristics in thicker oxide cells.

ACKNOWLEDGMENT

The authors would like to thank the Technology Development Center of MXIC for providing technical support.

REFERENCES

[1] B. Eitan, P. Pavan, I. Bloom, E. Aloni, A. Frommer, and D. Finzi, "NROM: A novel localized trapping, 2-bit nonvolatile memory cell," *IEEE Electron Device Lett.*, vol. 21, no. 11, pp. 543–545, Nov. 2000.

[2] M. H. White, "On the Go with SONOS," *IEEE Circuits Devices Mag.*, vol. 16, no. 4, pp. 22–31, Jul. 2000.

[3] M. K. Cho and D. M. Kim, "High performance SONOS memory cells free of drain turn-on and over-erase: Compatibility issue with current flash technology," *IEEE Electron Device Lett.*, vol. 21, no. 8, pp. 399–401, Aug. 2000.

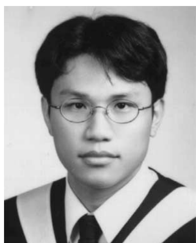
[4] T. Wang, W. J. Tsai, S. H. Gu, C. T. Chan, C. C. Yeh, N. K. Zous, T. C. Lu, S. Pan, and C.-Y. Lu, "Reliability models of data retention and read-disturb in 2-bit nitride storage flash memory cells (Invited Paper)," in *IEDM Tech. Dig.*, 2003, pp. 169–172.

[5] M. H. White, D. A. Adams, J. R. Murray, S. Wrazien, Y. Zhao, Y. Wang, B. Khan, W. Miller, and R. Mehrotra, "Characterization of scaled SONOS EEPROM memory devices for space and military systems," in *Proc. Symp. NVM Technol.*, 2004, pp. 51–59.

[6] T. H. Kim, J. S. Sim, J. D. Lee, H. C. Shim, and B.-G. Park, "Charge decay characteristics of silicon-oxide-nitride-oxide-silicon structure at elevated temperatures and extraction the nitride trap density distribution," *Appl. Phys. Lett.*, vol. 85, no. 4, pp. 660–662, Jul. 2004.

[7] H. Bachhofer, H. Reisinger, E. Bertagnolli, and H. von Philipsborn, "Transient conduction in multidielectric silicon-oxide-nitride-oxide semiconductor structures," *J. Appl. Phys.*, vol. 89, no. 5, pp. 2791–2800, Mar. 2001.

[8] L. Lundkvist, I. Lundstorm, and C. Svensson, "Discharge of MNOS structures," *Solid State Electron.*, vol. 16, no. 7, pp. 811–818, Jul. 1973.
[9] L. Lundkvist, C. Svensson, and B. Hansson, "Discharge of MNOS structures at elevated temperatures," *Solid State Electron.*, vol. 19, no. 3, pp. 221–227, Mar. 1976.
[10] P. J. McWhorter, S. L. Miller, and T. A. Dellin, "Modeling the memory retention characteristics of silicon-nitride-oxide-silicon nonvolatile transistors in a varying thermal environment," *J. Appl. Phys.*, vol. 68, no. 4, pp. 1902–1909, Aug. 1990.
[11] K. Lehovec and A. Fedotowsky, "Charge retention of MNOS devices limited by Frenkel–Poole detrapping," *Appl. Phys. Lett.*, vol. 32, no. 5, pp. 335–338, Mar. 1978.
[12] R. A. Williams and M. M. E. Beguwala, "The effect of electrical conduction of Si_3N_4 on the discharge of MNOS memory transistors," *IEEE Trans. Electron Devices*, vol. ED-25, no. 8, pp. 1019–1023, Aug. 1978.
[13] Y. Yang and M. H. White, "Charge retention of scaled SONOS nonvolatile memory devices at elevated temperatures," *Solid State Electron.*, vol. 44, no. 6, pp. 949–958, Jun. 2000.
[14] Y. Wang and M. H. White, "An analytical retention model for SONOS nonvolatile memory devices in the excess electron state," *Solid State Electron.*, vol. 49, no. 1, pp. 97–107, Jan. 2005.
[15] G. L. Chindalore, C. T. Swift, and D. Burnett, "A new combination-erase technique for erasing nitride based (SONOS) nonvolatile memories," *IEEE Electron Device Lett.*, vol. 24, no. 4, pp. 257–259, Apr. 2003.
[16] M. Specht, U. Dorda, L. Dreeskornfeld, J. Kretz, F. Hofmann, M. Städele, R. J. Luyken, W. Rösner, H. Reisinger, E. Landgraf, T. Schulz, J. Hartwich, R. Kömmling, and L. Risch, "20 nm tri-gate SONOS memory cells with multi-level operation," in *VLSI Symp. Tech. Dig.*, 2004, pp. 244–245.
[17] T. Wang, N. K. Zous, and C. C. Yeh, "Role of positive trapped charge in stress-induced leakage current for Flash EEPROM devices," *IEEE Trans. Electron Devices*, vol. 49, no. 11, pp. 1910–1916, Nov. 2002.
[18] V. Vasudevan and J. Vasi, "A simulation of multiple trapping model for continuous time random walk transport," *J. Appl. Phys.*, vol. 74, no. 5, pp. 3224–3230, Sep. 1993.
[19] D. Debuf, "General theory of carrier lifetime in semiconductors with multiple localized states," *J. Appl. Phys.*, vol. 96, no. 11, pp. 6454–6469, Dec. 2004.
[20] K. A. Nasyrov, V. A. Gritsenko, M. K. Kim, H. S. Chae, S. D. Chae, W. I. Ryu, J. H. Sok, J.-W. Lee, and B. M. Kim, "Charge transport mechanism in metal-nitride-oxide-silicon structures," *IEEE Electron Device Lett.*, vol. 23, no. 6, pp. 336–338, Jun. 2002.
[21] B. De Salvo, G. Ghibaudo, G. Pananakakis, P. Masson, T. Baron, N. Buffet, A. Fernandes, and B. Guillaumot, "Experimental and theoretical investigation of nano-crystal and nitride-trap memory devices," *IEEE Trans. Electron Devices*, vol. 48, no. 8, pp. 1789–1799, Aug. 2001.
[22] S. Kamohara, D. Park, and C. Hu, "Deep-trap SILC (stress induced leakage current) model for nominal and weak oxides," in *Proc. Int. Rel. Phys. Symp.*, 1998, pp. 57–61.
[23] M. Silver and L. Cohen, "Monte Carlo simulation of anomalous transit-time dispersion of amorphous solids," *Phys. Rev. B, Condens. Matter*, vol. 15, no. 6, pp. 3276–3278, Mar. 1977.
[24] C. Main, S. Reynolds, and R. Brüggemann, "Decay from steady-state photocurrent in amorphous semiconductors," *Phys. Stat. Sol. (C)*, vol. 1, no. 5, pp. 1194–1207, 2004.
[25] C. Main, J. Berkin, and A. Merazga, *New Physical Problems in Electronic Materials*. London, U.K.: Academic, 1992.
[26] T. R. Oldham, A. J. Lelis, and F. B. Mclean, "Spatial dependence of trapped holes determined from tunneling analysis and measured annealing," *IEEE Trans. Nucl. Sci.*, vol. NS-33, no. 8, pp. 1203–1209, Dec. 1986.
[27] K. I. Lundström and C. M. Svensson, "Properties of MNOS structures," *IEEE Trans. Electron Devices*, vol. ED-19, no. 6, pp. 826–836, Jun. 1972.



Shaw-Hung Gu was born in Taipei, Taiwan, R.O.C., on August 11, 1977. He received the B.S. degree from Chang-Gung University, Taoyuan, Taiwan, in 1999 and the M.S. degree in electronics engineering from the National Chiao-Tung University, Hsinchu, Taiwan, in 2001, where he is currently working toward the Ph.D. degree.

His research interest focuses on the physics and characterization of nonvolatile memory, especially in nitride-based storage memory and thin-oxide reliability.

Chih-Wei Hsu was born in Taipei, Taiwan, R.O.C., on August 17, 1981. He received the B.S. degree from Chang-Gung University, Taoyuan, Taiwan, in 2003 and the M.S. degree in electronics engineering from the National Chiao-Tung University, Hsinchu, Taiwan, in 2005.

His research interest focuses on nitride storage Flash memory, especially in numerical simulation of the retention behavior.

Tahui Wang (S'85–M'86–SM'94) was born in Tao-Yuan, Taiwan, R.O.C., on May 3, 1958. He received the B.S.E.E. degree from National Taiwan University, Taipei, Taiwan, in 1980 and the Ph.D. degree in electrical engineering from the University of Illinois, Urbana–Champaign, in 1985.

From 1985 to 1987, he was with Hewlett–Packard Laboratories, Palo Alto, CA, where he was engaged in the development of GaAs HEMT devices and circuits. Since 1987, he has been with the Department of Electronics Engineering, National Chiao-Tung University, Hsinchu, Taiwan, where he is currently a Professor. His research interests include hot-carrier phenomena characterization and reliability physics in very large-scale integration (VLSI) devices, RF CMOS devices, and nonvolatile semiconductor devices.

Dr. Wang is the recipient of the Best Teacher Award by the Ministry of Education, China. He has served as technical committee member of many international conferences, among them IEDM, IRPS, and VLSI-TSA. He was an invited speaker of 2003 IEDM on the topic of nitride Flash reliability.

Wen-Pin Lu was born in I-Lan, Taiwan, R.O.C., on December 20, 1967. He received the B.S. degree in electronics engineering from National Chiao-Tung University, Hsinchu, Taiwan, in 1990 and the M.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1992.

He joined Macronix International Company, Ltd., Hsinchu, in 1994, as a Device Engineer. From 1994 to 1999, he has worked on device analysis of nonvolatile memory, especially in floating-gate Flash memory. Since 2000, he has been engaged in the development of PACAND Flash-memory technology, and has accomplished 0.18 and 0.15 μm of delivering. He is presently responsible for the nitrite-based NBit technology integration at Macronix.

Yen-Hui Joseph Ku received the B.S. degree in electrical engineering from National Cheng-Kung University, Tainan, Taiwan, R.O.C., in 1979 and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Texas at Austin, in 1983 and 1988, respectively. His research was on the self-aligned silicide and shallow-junction formation for deep submicrometer device application.

In 1987, he joined Rapro Technology, Fremont, CA, as a Cofounder with the responsibility on RTPCVD reactor design and process development. From 1991 to 1992, he was with Paradigm Technology, San Jose, CA, where he worked on 0.35-μm 4-Mb SRAM technology development. In 1992, he joined LSI Logic, Santa Clara, CA, where he worked on the advanced logic technology development and served as Integration Manager of the R&D Division until 1997. He later joined the Technology Development Division of the Macronix International Company, Ltd., Hsinchu, Taiwan, R.O.C. He currently leads the Technology Development Center and is responsible for the development of advanced nonvolatile memory and embedded SOC technologies. He has published more than 30 papers in technical journals and conferences.

Chih-Yuan Lu 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 has been a Professor with National Chiao-Tung University and with AT&T Bell Laboratories from 1984 to 1989. He later joined Electronics Research and Services Organization (ERSO)/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 becomes the Cofounder of Vanguard International Semiconductor Corporation, which is a spin-off memory IC Company from ITRI's Submicron Project. He was the Vice President of Operation, Vice President of R&D, and later, President from 1994 to 1999. He is currently the Chairman and Chief Executive Officer of Ardentec Corporation, a very large-scale integration testing service company. He also serves Macronix International (MXIC) as a Senior Vice President/CTO. He led MXIC's technology development team to successfully achieve the state-of-the-art nonvolatile memory technology and is currently responsible for MXIC's overall memory operation. He has published more than 100 papers and is the holder of 123 international patents.

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